Linear Algebra and its Applications 433 (2010) 1491–1512

Contents lists available at ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

Block-oriented *J*-Jacobi methods for Hermitian matrices☆

Vjeran Hari [a,*], Sanja Singer [b], Saša Singer [a]

[a] *Department of Mathematics, University of Zagreb, P.O. Box 335, 10002 Zagreb, Croatia*
[b] *Faculty of Mechanical Engineering and Naval Architecture, University of Zagreb, I. Lučića 5, 10000 Zagreb, Croatia*

ARTICLE INFO

Article history:
Received 13 August 2008
Accepted 7 June 2010
Available online 3 August 2010

Submitted by V. Mehrmann

AMS classification:
65F15
65Y20
46C20

Keywords:
J-Jacobi algorithms
Block-oriented pivot strategies
Convergence
Efficiency

ABSTRACT

The paper describes a way how one-sided Jacobi-type algorithm of Veselić for computing the hyperbolic singular value decomposition of rectangular matrices can be modified to work with blocks. The proposed modification preserves the relative accuracy property of the original algorithm and essentially improves its performance. Special attention is devoted to proving the global convergence of the method under some important classes of block-oriented pivot strategies. As numerical tests indicate, the block-oriented *J*-Jacobi methods combined with the Hermitian indefinite factorization, become efficient and accurate eigensolvers for Hermitian indefinite matrices.

1. Introduction

In his pioneering paper [33], Veselić proposed a Jacobi-type algorithm for diagonalizing the definite matrix pairs (H, J), where H is symmetric and J is a diagonal matrix of signs. After a suitable preprocessing, the algorithm works with one matrix, using J-orthogonal congruences. The global convergence of the algorithm is proved in [33], while the asymptotic quadratic convergence is proved in [11,27].

The simplest and most natural application of the algorithm is to use it in the compound method for accurate computation of the eigenvalues and eigenvectors of an indefinite symmetric or Hermitian

☆ This work was supported by Grants 037-0372783-3042 and 037-1193086-2771 by Ministry of Science, Education and Sports, Croatia.

* Corresponding author.
E-mail addresses: hari@math.hr (V. Hari), ssinger@math.hr (S. Singer), singer@math.hr (S. Singer).

0024-3795/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2010.06.032

matrix of order n. The first phase of the method computes the indefinite factorization of the initial matrix H, by a variant of the Bunch–Parlett factorization (see [2,4–8]), $\widehat{P}H\widehat{P}^T = MDM^*$. Here $\widehat{P}$ is a permutation matrix, M is a lower triangular and $D = D^*$ is a block-diagonal matrix with diagonal blocks of order 1 or 2. An additional diagonalization of the diagonal blocks of D and an appropriate scaling of the columns of M yield

$$\widehat{P}H\widehat{P}^T = GJG^*, \qquad J = \mathrm{diag}(j_{11}, \ldots, j_{nn}), \tag{1.1}$$

where G is nonsingular lower block-triangular with diagonal blocks of order one or two, and $j_{ii} \in \{-1, 1, 0\}$ for $1 \leqslant i \leqslant n$. If H is nonsingular, J does not contain any zero along the diagonal.

If H is a singular $m \times m$ Hermitian matrix of rank n, then J contains $m - n$ zeros on the diagonal. Using the similarity transformation with a suitable permutation matrix, one can obtain J in the form $J = \mathrm{diag}(J_1, 0)$, where J_1 is of rank n. In this case, the relation (1.1) can be transformed into $H = G_1 J_1 G_1^*$, where G_1 is $m \times n$ matrix of full column rank and J_1 contains only ± 1 on the diagonal. Thus, we can assume that the relation (1.1) holds with nonsingular J, $\widehat{P} = I_m$ and full column rank $m \times n$ matrix G.

By pre-multiplying the equation $Hx = \lambda x$ with G^* and using (1.1), one obtains $G^*Gz = \lambda Jz$ with $z = JG^*x$. So, one has to solve the eigenproblem for the pair (G^*G, J), with positive definite G^*G. This is equivalent to solving the eigenproblem of the J-Hermitian matrix JG^*G. The eigenvalues of JG^*G and of the pair (G^*G, J) are exactly the non-zero eigenvalues of H.

In addition, one can apply the QR factorization with column pivoting to G,

$$G = QRP_1^T.$$

Then $G^*G = P_1 R^* R P_1^T$, so the pair (G^*G, J) can be transformed into $\left(R^*R, P_1^T J P_1\right)$. Note that $P_1^T J P_1$ is just another J. The advantage of this additional transformation is twofold. First, if $m > n$, then the matrix R has less rows than G and the later transformations will be cheaper. Second, the upper-triangular R has a special structure which implies that the rows of R can be well scaled if the columns of G can be well scaled. In particular, one can easily show that the condition number of $|R_r|$, is not larger than n times the condition number of $|R_c|$, where $R_r = \Delta_r R$ ($R_c = R\Delta_c$) and Δ_r (Δ_c) is so chosen that the rows of R_r (columns of R_c) have unit Euclidean norm (see [12, Proposition 3.2]). Usually the condition of R_r is considerably smaller than that of R_c. The assumption of small condition of R_r is important for proving the accuracy of simple and block Jacobi and J-Jacobi methods (see [25, Section 4.1], [19, Section 3.2], [23]).

The second phase of the method computes the eigenvalues and eigenvectors of the positive definite pair (G^*G, J) (or even better of (R^*R, J)). The most appropriate way to do it, is to compute the hyperbolic singular value decomposition (HSVD) of G (see [3]), by using the one-sided version of the Jacobi-type algorithm of Veselić [33] (see the relation (2.1)). We shall refer to it as J-Jacobi algorithm. This method has been proved to be relatively accurate [32]. In particular, if the condition number of $G\Delta$ is small for some nonsingular diagonal Δ, then the one-sided algorithm of Veselić will compute the hyperbolic singular values and vectors of G to high relative accuracy. This also ensures that the eigensystem of H will be computed accurately by that method.

If G is any $m \times n$ matrix with $m \geqslant n$, the HSVD of G with respect to J, has the form

$$G = U \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} V^*, \qquad \Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n), \quad \sigma_1 \geqslant \sigma_2 \geqslant \cdots \sigma_n \geqslant 0,$$

where U is $m \times m$ unitary, Σ is the matrix of hyperbolic singular values, and V is $n \times n$ J-unitary matrix, which satisfies $V^*JV = J$. If G is the factor of the Hermitian indefinite matrix H, $H = GJG^*$, then using the HSVD, we have

$$H = GJG^* = U \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} V^* J V \begin{bmatrix} \Sigma & 0 \end{bmatrix} U^* = U \begin{bmatrix} J\Sigma^2 & 0 \\ 0 & 0 \end{bmatrix} U^*.$$

Hence, the squares of the hyperbolic singular values of G are, up to the signs in J, the non-zero eigenvalues of H, and U is the corresponding eigenvector matrix.

Our tests have shown that the J-Jacobi method is more efficient provided that $J = \mathrm{diag}(I_\nu, -I_{n-\nu})$ where $1 \leqslant \nu \leqslant n - 1$. This can be achieved via the congruence transformation

$(G^*G, J) \mapsto (P_2^* G^* G P_2, P_2^* J P_2)$ with a suitable permutation matrix P_2. It has been observed that the J-Jacobi algorithm is on average 25–30% slower than the standard Jacobi algorithm. This obviously comes from using the hyperbolic transformations, which can temporarily divert the off-diagonal part of the current iterate from converging to zero. If the partition of G^*G respects the signs in J, more block-pivot submatrices of G^*G will be definite and consequently we can expect faster convergence.

Finally, note that post-multiplying G by P_2, does not change the minimum of the condition number of $G\Delta$ over all Δ, so we shall again have small condition of R_r. This shows that one should compute HSVD of R instead of G, as mentioned above.

Now, let us concentrate on the iterative part of the compound method for H. We know (see [19]) that one-sided Jacobi algorithms can be made more efficient by blocking. In the presence of two or more layers of memory with different speeds, many blocked algorithms show significant speedups. Even simple matrix multiplication performed by blocks (for example, implemented as `xGEMM` BLAS-3 subroutine) is several times faster than the standard multiplication by elements. A similar reasoning leads us to the construction of the block Jacobi-type algorithms (see [18,19,22]).

Two ways how to design the block J-Jacobi algorithms are described in Section 2. We consider the *block-oriented* algorithms and only mention the *full block* algorithms since they are considered in another paper [23]. During construction, special care is taken to assure the overall quadratic convergence of the algorithms. Otherwise, efficiency of the algorithms can be lost. We note that stopping of the one-sided (simple or block) algorithms is not yet quite satisfactorily solved.

The accuracy result from [32] cannot be used directly here, since any block method applies the block-transformations $\mathbf{V}^{(k)}$ (from the relation (2.6)) and not just the sequence of plane transformations to the iterated matrix. We shall not consider here the relative accuracy of the block-oriented method since the accuracy proof for the full block method from [23] holds also for the block-oriented method.

Our numerical experiments indicate, that for large enough dimension of H, our block-oriented algorithms can be 40% faster than their simple (non-blocked) counterparts. As expected, our tests also show that the proposed block algorithm inherits the relative accuracy from the simple one-sided J-Jacobi algorithm (cf. [9]). However, the major attention in the paper is devoted to the global convergence of the block-oriented algorithms.

The rest of the paper is organized as follows. In Section 2, we summarize the known facts for the simple algorithm of Veselić and we derive similar facts for the block-oriented two-sided algorithms. We also define three classes of block strategies. In Section 3, we consider the global convergence for the associated block-oriented algorithms. A note on the asymptotic convergence behavior of the algorithms is also given. In Section 4, we describe fine implementation details for the one-sided block-oriented algorithms. In Section 5, we present the results of numerical tests. It also includes accurate timing for the iterative part of the considered block algorithm. Finally, in Appendix we provide the lengthy and/or technical proofs of the results from Section 3.

2. Two-sided block J-Jacobi methods

Although, for accuracy and efficiency reasons, each J-Jacobi method should be implemented as one-sided algorithm, the global and the asymptotic convergence investigation uses its two-sided version. In contrast to accuracy issues, the convergence considerations presume infinite arithmetic and therefore the convergence of one-sided method means the convergence of the corresponding two-sided method.

If we write $A = G^*G$, the two-sided and the one-sided versions of the J-Jacobi method are described by the following two processes

$$\begin{aligned} A^{(k+1)} &= [V^{(k)}]^* A^{(k)} V^{(k)}, \qquad k \geqslant 0, \qquad A^{(0)} = A, \\ G^{(k+1)} &= G^{(k)} V^{(k)}, \qquad\quad k \geqslant 0, \qquad G^{(0)} = G, \end{aligned} \tag{2.1}$$

where $V^{(k)}$ are plane J-unitary matrices. The method is said to be globally convergent if for each initial A, the generated sequence $(A^{(k)}, k \geqslant 0)$ tends to some diagonal matrix. We shall measure convergence by the *off-norm* of $A^{(k)}$. The off-nom is defined by

$$\mathrm{Off}(X) = \|X - \mathrm{diag}(X)\|, \tag{2.2}$$

where $\mathrm{diag}(X) = \mathrm{diag}(x_{11}, \ldots, x_{nn})$ is the diagonal part of $X = (x_{ij})$ and

$$\|X\| = [\mathrm{tr}(X^*X)]^{1/2} = \left(\sum_{i=1}^{n} \sum_{j=1}^{n} |x_{ij}|^2 \right)^{1/2}$$

is the Frobenius norm of X. Here $\mathrm{tr}(X^*X)$ is the trace (i.e., the sum of the diagonal elements) of X^*X.

The method is said to be (asymptotically) quadratically convergent if the off-norm of $A^{(k)}$ tends to zero quadratically per sweep, which consists of $n(n-1)/2$ consecutive steps.

2.1. Using blocks

We shall describe ways how to make the J-Jacobi a BLAS 3 algorithm. Essentially, there are two ways how to do that: using some block-oriented cyclic or quasi-cyclic pivot strategy or using a full block algorithm (which diagonalizes the block-pivot submatrix at each step) under an appropriate pivot strategy. In this paper we concentrate on the first way.

A few words about notation. Since we deal with matrices which carry two levels of block-matrix partitions, one inherited from J, the other assumed in accordance with the cache memory capacity, we shall use boldface $\mathbf{J}$ for the starting J and $\mathbf{A}$ for the starting iterated matrix.

Thus, we start our consideration with the pair $(\mathbf{A}, \mathbf{J})$, where $\mathbf{A}$ is positive definite and $\mathbf{J} = \mathrm{diag}(I_\nu, -I_{n-\nu})$, $1 \leqslant \nu \leqslant n-1$. Typically, $\mathbf{A} = G^*G$, $G = RP$, where $P = P_1P_2$ is permutation. The permutation P_1 and the upper-triangular R come from the QR factorization with column pivoting of G, while the permutation P_2 transforms J into $\mathbf{J}$. In addition, P_2 is chosen to make the diagonal elements of $\mathbf{JA}$ non-increasingly ordered.

If we partition the matrix $\mathbf{A} = (a_{rt})$ in accordance with the *initial partition* of $\mathbf{J}$, we can write

$$\mathbf{A} = \left[\begin{array}{c|c} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \hline \mathbf{A}_{12}^* & \mathbf{A}_{22} \end{array}\right], \quad \mathbf{J} = \left[\begin{array}{c|c} I_\nu & \\ \hline & -I_{n-\nu} \end{array}\right], \quad \begin{array}{l} a_{11} \geqslant a_{22} \geqslant \cdots \geqslant a_{\nu\nu}, \\ a_{\nu+1,\nu+1} \leqslant \cdots \leqslant a_{nn}. \end{array} \tag{2.3}$$

Here $\mathbf{A}_{11}$ is $\nu \times \nu$ positive definite matrix whose diagonal elements are non-increasingly ordered. The diagonals of $\mathbf{A}_{22}$ are ordered non-decreasingly. This assumption is attractive for two reasons: it makes theoretical analysis simpler and, as numerical tests indicate, J-Jacobi methods converge faster if this property is present during iteration.

Let us explain that. Let $\mathbf{C}$ be a $\mathbf{J}$-unitary matrix which diagonalizes $\mathbf{A}$. That is, $\mathbf{C}^*\mathbf{JC} = \mathbf{J}$, $\mathbf{C}^*\mathbf{AC} = \mathbf{\Lambda}$, where $\mathbf{\Lambda}$ is diagonal. Then

$$\mathbf{J\Lambda} = \mathbf{JC}^*\mathbf{AC} = \mathbf{JC}^*(\mathbf{JJ})\mathbf{AC} = (\mathbf{JC}^*\mathbf{J})(\mathbf{JA})\mathbf{C} = \mathbf{C}^{-1}\mathbf{JAC}.$$

If the eigenvalues of $\mathbf{JA}$ are ordered non-increasingly, the perturbation analysis will be simpler if the same is assumed for the diagonal elements of $\mathbf{JA}$, and this reduces to the assumption in (2.3). Let $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Then the eigenvalues of $\mathbf{JA}$ are $\lambda_1 \geqslant \cdots \geqslant \lambda_\nu > -\lambda_{\nu+1} \geqslant \cdots \geqslant -\lambda_n$. In [33], Veselić has shown that the gap between the positive and the negative part of the spectrum, $\delta = \lambda_\nu + \lambda_{\nu+1}$, satisfies the inequality $a_{rr} + a_{ss} > \delta$, whenever $1 \leqslant r \leqslant \nu < s \leqslant n$. Since δ and $\mathbf{J}$ are invariant under $\mathbf{J}$-unitary transformations, this property will hold at any step of the method.

Since $\mathbf{A}$ is positive definite, one can show that the Hermitian matrix $\mathbf{A} - \mu\mathbf{J}$ is positive definite if and only if $\mu \in \langle -\lambda_{\nu+1}, \lambda_\nu \rangle$. The classical perturbation theorem for Hermitian matrices implies that $\mathbf{A} - \mu\mathbf{J}$ will be positive definite for any $\mu \in \langle -\sigma_{\min}(\mathbf{A}), \sigma_{\min}(\mathbf{A}) \rangle$. Hence, we must have $2\sigma_{\min}(\mathbf{A}) \leqslant \delta$. Here, $\sigma_{\min}(\mathbf{A})$ is the smallest singular value (which is also the eigenvalue) of $\mathbf{A}$.

Each block $\mathbf{J}$-Jacobi method is defined by some "block pivot strategy" which selects the off-diagonal blocks, one at a time. Therefore, a block-matrix partition must be given. We call it the *basic (block) partition* and denote it in the following way

$$\mathbf{A}_{11} = \begin{bmatrix} A_{11} & \cdots & A_{1p} \\ \vdots & \ddots & \vdots \\ A_{1p}^* & \cdots & A_{pp} \end{bmatrix}, \qquad \mathbf{A}_{12} = \begin{bmatrix} A_{1,p+1} & \cdots & A_{1,p+q} \\ \vdots & \ddots & \vdots \\ A_{p,p+1} & \cdots & A_{p,p+q} \end{bmatrix},$$

$$\mathbf{A}_{12}^{*}=\begin{bmatrix} A_{1,p+1}^{*} & \cdots & A_{p,p+1}^{*} \\ \vdots & \ddots & \vdots \\ A_{1,p+q}^{*} & \cdots & A_{p,p+q}^{*} \end{bmatrix}, \qquad \mathbf{A}_{22}=\begin{bmatrix} A_{p+1,p+1} & \cdots & A_{p+1,p+q} \\ \vdots & \ddots & \vdots \\ A_{p+1,p+q}^{*} & \cdots & A_{p+q,p+q} \end{bmatrix}. \tag{2.4}$$

Here, each diagonal block A_{ii} is of order n_i, $1 \leqslant i \leqslant p+q$. The same partition applies to $\mathbf{J}$, so that

$$\mathbf{J} = \operatorname{diag}(J_{11}, \ldots, J_{pp}, \ldots, J_{p+q,p+q}).$$

Obviously, the two partitions have to satisfy $n_1 + \cdots + n_p = \nu$.

At one (block-) step, a two-sided block $\mathbf{J}$-Jacobi method will either annihilate the pivot off-diagonal blocks A_{ij} and A_{ji} or will just reduce their Frobenius norm. Also, if the pivot diagonal block A_{ii} is of order $n_i > 1$, then the method will either reduce its off-norm to zero or will just reduce it by some amount.

Therefore, we describe a (two-sided) block $\mathbf{J}$-Jacobi method for solving the eigenproblem of the pair $(\mathbf{A}, \mathbf{J})$, as an iterative process of the form

$$\mathbf{A}^{(k+1)} = [\mathbf{V}^{(k)}]^{*}\mathbf{A}^{(k)}\mathbf{V}^{(k)}, \quad k \geqslant 0, \quad \mathbf{A}^{(0)} = \mathbf{A}, \tag{2.5}$$

where each $\mathbf{V}^{(k)}$ is $\mathbf{J}$-unitary, i.e., $[\mathbf{V}^{(k)}]^{*}\mathbf{J}\mathbf{V}^{(k)} = \mathbf{J}$, and has one of the following forms

$$\mathbf{V}^{(k)} = \begin{bmatrix} I & & & & \\ & V_{ii}^{(k)} & & V_{ij}^{(k)} & \\ & & I & & \\ & V_{ji}^{(k)} & & V_{jj}^{(k)} & \\ & & & & I \end{bmatrix} \begin{matrix} \\ \} \ n_i \\ \\ \} \ n_j \\ \\ \end{matrix} \quad \text{if} \quad i < j,$$

$$\mathbf{V}^{(k)} = \begin{bmatrix} I & & \\ & V_{ii}^{(k)} & \\ & & I \end{bmatrix} \quad \text{if} \quad i = j, \quad n_i > 1. \tag{2.6}$$

Here, $i = i(k), j = j(k), i \leqslant j$, are the *pivot indices* and (i, j) is the *pivot pair*. *Pivot strategy* is a way how the pivot pairs are selected. Since i and j are subscripts of blocks, we shall frequently use the phrases: *block pivot indices*, *block pivot pair* and *block pivot strategy* or shorter *block strategy*.

The *block pivot submatrix* (or shorter, the *pivot submatrix*) of $\mathbf{V}^{(k)}$ is the matrix

$$\widehat{\mathbf{V}}^{(k)} = \begin{cases} \begin{bmatrix} V_{ii}^{(k)} & V_{ij}^{(k)} \\ V_{ji}^{(k)} & V_{jj}^{(k)} \end{bmatrix}, & i < j, \\ V_{ii}^{(k)}, & i = j, \ n_i > 1. \end{cases} \tag{2.7}$$

The matrix $\widehat{\mathbf{V}}^{(k)}$ is $\widehat{\mathbf{J}}$-unitary, where $\widehat{\mathbf{J}} = \operatorname{diag}(J_{ii}, J_{jj})$. In fact, $\widehat{\mathbf{V}}^{(k)}$ and hence $\mathbf{V}^{(k)}$ are unitary provided that $1 \leqslant i < j \leqslant p$ or $p+1 \leqslant i < j \leqslant p+q$ or $1 \leqslant i = j \leqslant p+q$ and $n_i > 1$.

2.2. One block step

Here we consider one block step of a block-oriented method. For simplicity, we denote the current matrix $\mathbf{A}^{(k)}$ by $\mathbf{A}$, the transformed matrix $\mathbf{A}^{(k+1)}$ by $\mathbf{A}'$ and the $\mathbf{J}$-unitary transformation matrix by $\mathbf{V}$. We call the principal submatrix of $\mathbf{A}$, which is transformed by both, the left-hand and the right-hand transformation, *pivot submatrix* of $\mathbf{A}$ and denote it by $\widehat{\mathbf{A}}$. This is in accordance with the notation of $\widehat{\mathbf{V}}^{(k)}$ from (2.7). Note that

$$\mathbf{J}\mathbf{A}' = \mathbf{J}\mathbf{V}^{*}\mathbf{A}\mathbf{V} = \mathbf{J}\mathbf{V}^{*}\mathbf{J}(\mathbf{J}\mathbf{A})\mathbf{V} = \mathbf{V}^{-1}(\mathbf{J}\mathbf{A})\mathbf{V},$$

hence $(\mathbf{J}\mathbf{A}')^2 = \mathbf{V}^{-1}(\mathbf{J}\mathbf{A})^2\mathbf{V}$. Therefore, (cf. [33])

$$\operatorname{tr}(\mathbf{J}\mathbf{A}') = \operatorname{tr}(\mathbf{J}\mathbf{A}), \qquad \operatorname{tr}((\mathbf{J}\mathbf{A}')^2) = \operatorname{tr}((\mathbf{J}\mathbf{A})^2)$$

and since

$$\mathbf{A} = \left[\begin{array}{c|c} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \hline \mathbf{A}_{12}^* & \mathbf{A}_{22} \end{array}\right], \qquad \mathbf{A}' = \left[\begin{array}{c|c} \mathbf{A}'_{11} & \mathbf{A}'_{12} \\ \hline [\mathbf{A}'_{12}]^* & \mathbf{A}'_{22} \end{array}\right],$$

we have (see [33])

$$\begin{aligned} &\mathrm{tr}(\mathbf{A}'_{11}) - \mathrm{tr}(\mathbf{A}'_{22}) = \mathrm{tr}(\mathbf{A}_{11}) - \mathrm{tr}(\mathbf{A}_{22}) \\ &2\|\mathbf{A}'_{12}\|^2 - \|\mathbf{A}'_{11}\|^2 - \|\mathbf{A}'_{22}\|^2 = 2\|\mathbf{A}_{12}\|^2 - \|\mathbf{A}_{11}\|^2 - \|\mathbf{A}_{22}\|^2. \end{aligned}$$

On the level of pivot submatrices, we have $\widehat{\mathbf{A}}' = \widehat{\mathbf{V}}^* \widehat{\mathbf{A}} \widehat{\mathbf{V}}$, that is

$$\begin{bmatrix} A'_{ii} & A'_{ij} \\ [A'_{ij}]^* & A'_{jj} \end{bmatrix} = \begin{bmatrix} V_{ii}^* & V_{ji}^* \\ V_{ij}^* & V_{jj}^* \end{bmatrix} \begin{bmatrix} A_{ii} & A_{ij} \\ A_{ij}^* & A_{jj} \end{bmatrix} \begin{bmatrix} V_{ii} & V_{ij} \\ V_{ji} & V_{jj} \end{bmatrix} \quad \text{if} \quad i < j \tag{2.8}$$

and

$$A'_{ii} = V_{ii}^* A_{ii} V_{ii} \quad \text{if} \quad i = j. \tag{2.9}$$

The purpose of one step is to make $\mathbf{A}'$ more diagonal than $\mathbf{A}$. To this end one typically makes $\widehat{\mathbf{A}}'$ more diagonal than $\widehat{\mathbf{A}}$. As noted earlier, we distinguish two cases:

(i) the norm $\|A_{ij}\|$ is reduced, and
(ii) the off-diagonal block A_{ij} is annihilated.

In the first case we speak of a *block-oriented Jacobi-type method*. In the second case we speak of a *proper* or *full block method*. Thus, the proper block steps are part of the proper block methods, and (general) block steps are used in block-oriented methods. At one proper block step, we construct the transformation which annihilates A_{ij}; so at micro-level we have to use some method for this purpose. At one block step of a block-oriented method, we sequentially annihilate the elements of A_{ij} (and obviously of A_{ji}) in some way, typically in the row-or column-wise fashion, so in general, we perform easier job. For the case $\widehat{\mathbf{A}} = A_{ii}$ this means that in the case (i) we perform one sweep (or quasi-sweep) of unitary Jacobi steps on A_{ii} and in the case (ii) we diagonalize A_{ii}.

In (2.8), we refer to $\widehat{\mathbf{A}}$ as the pivot submatrix and to A_{ij} as the *pivot block* of $\mathbf{A}$. These two terms coincide in the relation (2.9) since there $\widehat{\mathbf{A}} = A_{ii}$.

2.3. One block step of a block-oriented method

Let us assume that the block A_{ij} from (2.8) is "operated" in such a way that $n_i \cdot n_j$ ordinary (simple) steps of the **J**-Jacobi method are applied to the current iterate $\mathbf{A}$, following the column-cyclic ordering within A_{ij}. We first consider a simple step which annihilates one matrix element.

Let a_{rs} be the pivot element. Then the simple **J**-unitary transformation has the form $\mathbf{A}' = V^* \mathbf{A} V$ and on the level of 2×2 pivot submatrices, we have

$$\widehat{A}' = \widehat{V}^* \widehat{A} \widehat{V}.$$

We can assume that $\widehat{V} = \widehat{\Phi}\widehat{R}$, where $\Phi = \Phi(\phi)$ is unitary and diagonal, while $R = R(\theta)$ is a real elementary plane matrix, orthogonal or $\widehat{J}$-orthogonal, where $\widehat{J}$ is the appropriate 2×2 submatrix of **J**. The role of Φ is to transform the Hermitian positive definite $\widehat{A}$ into a symmetric positive definite and non-negative $\widehat{\Phi}^* \widehat{A} \widehat{\Phi}$,

$$\widehat{\Phi}^* \widehat{A} \widehat{\Phi} = \begin{bmatrix} a_{rr} & |a_{rs}| \\ |a_{rs}| & a_{ss} \end{bmatrix}.$$

The task can be performed by the matrix $\widehat{\Phi} = \mathrm{diag}(e^{\iota\phi}, 1)$ or $\widehat{\Phi} = \mathrm{diag}(1, e^{-\iota\phi})$, assuming $\phi = \mathrm{Arg}(a_{rs})$. The second component $\widehat{R}$ can be defined as in [33]:

$$\widehat{R} = \begin{cases} \begin{bmatrix} \cosh\theta & \sinh\theta \\ \sinh\theta & \cosh\theta \end{bmatrix}, & \tanh 2\theta = \frac{-2|a_{rs}|}{a_{rr}+a_{ss}}, \quad \text{if } 1 \leqslant r \leqslant \nu < s \leqslant n \\ \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, & \tan 2\theta = \frac{2|a_{rs}|}{a_{rr}-a_{ss}}, \quad \text{elsewhere.} \end{cases} \tag{2.10}$$

Let us suppose first that $1 \leqslant r \leqslant \nu < s \leqslant n$ holds. Then we have (see [33,27])

$$a'_{rr} = a_{rr} + \tanh\theta \cdot |a_{rs}|, \qquad a'_{ss} = a_{ss} + \tanh\theta \cdot |a_{rs}|. \tag{2.11}$$

Hence, from (2.10) and (2.11) it follows

$$a'_{rr} + a'_{ss} = a_{rr} + a_{ss} + 2\tanh\theta \cdot |a_{rs}| = (a_{rr} + a_{ss})(1 - \tanh\theta\tanh 2\theta) = (a_{rr} + a_{ss})\kappa,$$

where

$$\kappa = 1 - \tanh\theta\tanh 2\theta = \sqrt{1 - \tanh^2 2\theta},$$

and

$$\begin{aligned}\operatorname{tr}(\mathbf{A}') - \operatorname{tr}(\mathbf{A}) &= a'_{rr} + a'_{ss} - (a_{rr} + a_{ss}) = -(a_{rr} + a_{ss})\tanh\theta\tanh 2\theta = (\kappa - 1)(a_{rr} + a_{ss}) \\ &= \frac{\kappa^2 - 1}{\kappa + 1}(a_{rr} + a_{ss}) = \frac{-\tanh^2 2\theta}{\kappa + 1}(a_{rr} + a_{ss}) = \frac{-4|a_{rs}|^2}{a_{rr} + a_{ss}} \cdot \frac{1}{1 + \kappa}.\end{aligned}$$

This implies

$$\tanh\theta\tanh 2\theta = \frac{\operatorname{tr}(\mathbf{A}) - \operatorname{tr}(\mathbf{A}')}{a_{rr} + a_{ss}} \leqslant \frac{\operatorname{tr}(\mathbf{A}) - \operatorname{tr}(\mathbf{A}')}{\delta}. \tag{2.12}$$

Thus, during "hyperbolic steps", the trace of $\mathbf{A}$ cannot increase, and it decreases if and only if $a_{rs} \neq 0$. If $\mathbf{A}$ is real symmetric, then $\phi = 0$ and $|a_{rs}|$ in (2.10) and (2.11) has to be replaced by a_{rs}.

If $1 \leqslant r \leqslant \nu < s \leqslant n$ does not hold, then the simple unitary Jacobi step which annihilates a_{rs} is applied. It is known that

$$a'_{rr} = a_{rr} + \tan\theta \cdot |a_{rs}|, \qquad a'_{ss} = a_{ss} - \tan\theta \cdot |a_{rs}|, \tag{2.13}$$

whence $a'_{rr} + a'_{ss} = a_{rr} + a_{ss}$ implying $\operatorname{tr}(\mathbf{A}') = \operatorname{tr}(\mathbf{A})$. Thus, in this case $\operatorname{tr}(\mathbf{A})$ remains invariant. It is also known that

$$\operatorname{Off}^2(\mathbf{A}) - \operatorname{Off}^2(\mathbf{A}') = 2|a_{rs}|^2,$$

where $\operatorname{Off}(\mathbf{A})$ is defined by (2.2). If $\mathbf{A}$ is real symmetric, then $\phi = 0$, and $|a_{rs}|$ in (2.10) and (2.13) has to be replaced by a_{rs}.

Now, let us consider the block steps of some block-oriented **J**-Jacobi method. We conclude that $\operatorname{tr}(\mathbf{A}^{(k)})$ cannot increase during one block step. Since the sequence $(\operatorname{tr}(\mathbf{A}^{(k)}), k \geqslant 0)$ is always bounded below by zero, it is convergent. Note that all $\mathbf{A}^{(k)} = [G^{(k)}]^* G^{(k)}$ are positive definite. Therefore, for all k,

$$\|\mathbf{A}^{(k)}\| = \sqrt{\sum_{r=1}^{n} \lambda_i^2(\mathbf{A}^{(k)})} \leqslant \sum_{r=1}^{n} |\lambda_i(\mathbf{A}^{(k)})| = \sum_{r=1}^{n} \lambda_i(\mathbf{A}^{(k)}) = \operatorname{tr}(\mathbf{A}^{(k)}) \leqslant \operatorname{tr}(\mathbf{A}), \tag{2.14}$$

$$\|G^{(k)}\|^2 = \operatorname{tr}(\mathbf{A}^{(k)}) \leqslant \operatorname{tr}(\mathbf{A}^{(k-1)}) = \|G^{(k-1)}\|^2 \leqslant \operatorname{tr}(\mathbf{A}) = \|G\|^2. \tag{2.15}$$

This implies that the sequence $(\mathbf{A}^{(k)}, k \geqslant 0)$ is contained in the ball of radius $\operatorname{tr}(\mathbf{A})$ and that the sequence $(\|G^{(k)}\|, k \geqslant 0)$ is non-increasing and convergent.

The relations (2.14) and (2.15) show that hyperbolic two-sided and one-sided transformations cannot essentially blow up the elements of $\mathbf{A}^{(k)}$ and $G^{(k)}$.

Let $t = 1, 2, \ldots$ count the simple steps and let its subsequence $\tilde{t}$ count the hyperbolic steps. From the relation (2.12), we can conclude that

$$\tanh\theta(\tilde{t})\tanh 2\theta(\tilde{t}) \to 0 \quad \text{as} \quad \tilde{t} \to \infty. \tag{2.16}$$

Since $|\tanh\theta| \leqslant |\tanh 2\theta| \leqslant 2|\tanh\theta|$, we have $\tanh^2 2\theta \leqslant 2\tanh\theta\tanh 2\theta$. Hence the relation (2.16) implies $\tanh 2\theta(\tilde{t}) \to 0$ as $\tilde{t} \to \infty$. Now,

$$\frac{2|a_{rs}|}{\delta} \leqslant \frac{2|a_{rs}|}{a_{rr} + a_{ss}} = \tanh 2\theta(\tilde{t}) \to 0 \quad \text{as} \quad \tilde{t} \to \infty, \tag{2.17}$$

implies that each pivot element a_{rs} of a hyperbolic step (from $\mathbf{A}_{12}$-block) tends to zero as $\tilde{t} \to \infty$. Furthermore, the hyperbolic angle $\theta(\tilde{t})$ tends to zero as $\tilde{t} \to \infty$.

2.4. Pivot strategies

Let $\mathbb{N}_0 = \{0, 1, 2, \ldots, \}$ and let the integers n, r be such that $n \geqslant r \geqslant 1, n \geqslant 2$. Let the partition $n = (n_1, \ldots, n_r), n_i \geqslant 1$ for $1 \leqslant i \leqslant r$, be given. We define

$$\mathbf{P}_r = \{(s,t) : 1 \leqslant s < t \leqslant r\} \cup \{(s,s) : n_s > 1,\ 1 \leqslant s \leqslant r\} \tag{2.18}$$

and $N_r = \mathrm{card}(\mathbf{P}_r)$, the cardinality of $\mathbf{P}_r$.

We first define the *block pivot strategies*. Once n, ν, p and q are given together with the partitions $n_1, \ldots, n_p$ and $n_{p+1}, \ldots, n_{p+q}$ of ν and $n - \nu$, respectively, we can define the block strategies as functions from $\mathbb{N}_0$ to $\mathbf{P}_{p+q}$, where $\mathbf{P}_{p+q}$ is given by (2.18) with $r = p + q$. Each block strategy $\mathcal{I}$ can be written as the function $\mathcal{I}(k) = (i(k), j(k))$, $k \geqslant 0$. If $\mathcal{I}$ is a periodic function, then $\mathcal{I}$ is called *periodic block strategy*.

Let $\mathcal{I}$ be a periodic block strategy with period M. If $M > N_{p+q}$ ($M = N_{p+q}$) and $\{\mathcal{I}(k) : k = 0, 1, \ldots, M-1\} = \mathbf{P}_{p+q}$, then $\mathcal{I}$ is called *quasi-cyclic* (*cyclic*) block strategy. All block strategies considered in this paper will be periodic, so the term *block strategy* will actually mean the periodic block pivot strategy.

2.4.1. Sequences and strategies

Let $\mathbf{S}$ be a subset of $\mathbf{P}_{p+q}$. By $\mathbf{O}(\mathbf{S})$ we denote the collection of all finite sequences made of the elements of $\mathbf{S}$. Suppose that $O \in \mathbf{O}(\mathbf{S})$. We assume that each element of $\mathbf{S}$ appears at least once in O. Otherwise $\mathbf{S}$ can be replaced by some of its proper subsets. Thus, each sequence from $\mathbf{O}(\mathbf{S})$ contains at least card($\mathbf{S}$) terms. A cyclic or a quasi-cyclic block strategy can be specified in the following way. For any sequence $O = (i_0, j_0), \ldots, (i_{M-1}, j_{M-1}) \in \mathbf{O}(\mathbf{P}_{p+q})$ the cyclic or quasi-cyclic block strategy $\mathcal{I}_O$ generated by O is given by

$$\mathcal{I}_O(k) \equiv (i(k), j(k)) = (i_t, j_t), \quad 0 \leqslant t \leqslant M-1, \quad k \geqslant 0,$$

provided that $k \equiv t$ (mod M). In other words, we have

$$\begin{array}{llll} \mathcal{I}_O(0) = (i_0, j_0), & \mathcal{I}_O(1) = (i_1, j_1), & \ldots, & \mathcal{I}_O(M-1) = (i_{M-1}, j_{M-1}), \\ \mathcal{I}_O(M) = (i_0, j_0), & \mathcal{I}_O(M+1) = (i_1, j_1), & \ldots, & \mathcal{I}_O(2M-1) = (i_{M-1}, j_{M-1}), \\ \mathcal{I}_O(2M) = (i_0, j_0), & \mathcal{I}_O(2M+1) = (i_1, j_1), & \ldots & \end{array}$$

When we speak of a block pivot strategy, or of selecting and transforming the whole blocks, we speak of the *macro-level* aspect of the method. The *micro-level* aspect will include all transformations, methods, pivot strategies,... which are performed or used within the blocks.

2.4.2. Simple strategies

Before we proceed with the macro-level considerations, let us recall some basic notions linked with the simple, i.e., non-block pivot strategies. Let $\mathbf{P} = \mathbf{P}_n$ so that $r = n$ and let $n_i = 1$for all $1 \leqslant i \leqslant r$. Then each pair of $\mathbf{P}$ addresses one element of the strictly upper-triangular part of $\mathbf{A}$. Let $\mathbf{S}$ be any subset of $\mathbf{P}$.

By $O_R(\mathbf{S})$ we denote the "row-wise ordering of $\mathbf{S}$", that is the sequence satisfying the following two conditions:

(a) each element $(i, j) \in \mathbf{S}$ appears exactly once in $O_R(\mathbf{S})$, and
(b) for any two terms (i_1, j_1) and (i_2, j_2) in $O_R(\mathbf{S})$, (i_1, j_1) precedes (i_2, j_2) if $i_1 < i_2$ or $i_1 = i_2$ and $j_1 < j_2$.

In an obvious manner we can define the "column-wise ordering of $\mathbf{S}$" denoted by $O_C(\mathbf{S})$. With an abuse of notation, we shall write $O_R(B)$, where B is a block (submatrix) in the strictly upper-triangular part of $\mathbf{A}$ or a principal submatrix of $\mathbf{A}$. In these situations, $O_R(B)$ will denote the row-wise ordering of the set of index pairs belonging to the elements of B which are located in the strictly upper-triangular part of $\mathbf{A}$. In the same way is defined $O_C(B)$. For example, $O_R(\mathbf{A})$ is the ordering $O_R(\mathbf{P})$.

2.4.3. Equivalent strategies

Let $\mathbf{S}$ be a subset of $\mathbf{P}$ and let $O = \{(i_r, j_r)\}_{r=0}^{s} \in \mathbf{O}(\mathbf{S})$. An *admissible transposition* on O is any transposition of two adjacent terms

$$(i_r, j_r), (i_{r+1}, j_{r+1}) \to (i_{r+1}, j_{r+1}), (i_r, j_r),$$

provided that the sets $\{i_r, j_r\}$ and $\{i_{r+1}, j_{r+1}\}$ are disjoint. The sequences $O, O' \in \mathbf{O}(\mathbf{S})$ are *equivalent* if one can be obtained from the other by a finite number of admissible transpositions. In this case we write $O \sim O'$.

Let $\mathcal{I}$ be a strategy with period M. By $O_{\mathcal{I}}$ we mean the sequence $\{\mathcal{I}(k)\}_{k=0}^{M-1}$. Now, let $\mathcal{I}$ and $\mathcal{I}'$ be two strategies with the same period M. The strategies $\mathcal{I}$ and $\mathcal{I}'$ are *equivalent*, if $O_{\mathcal{I}} \sim O_{\mathcal{I}'}$. In such a case we write $\mathcal{I} \sim \mathcal{I}'$.

If $\mathcal{I}$ and $\mathcal{I}'$ are equivalent, then the **J**-Jacobi processes defined by them yield the same matrices $\mathbf{A}^{(tM)}$, $t \geqslant 0$ (see [15–17,29]).

Let $\mathbf{S}_i$, $1 \leqslant i \leqslant \sigma$, be subsets of $\mathbf{P}$ and $O_i \in \mathbf{O}(\mathbf{S}_i)$, $1 \leqslant i \leqslant \sigma$. By $[O_1, O_2, \ldots, O_\sigma]$ or simply $O_1, O_2, \ldots, O_\sigma$ we mean the sequence which is obtained by the concatenation of the sequences $O_1, O_2, \ldots, O_\sigma$.

It is obvious that the notions of equivalent pivot strategies can be defined on macro-level, for block pivot strategies. To do that, we just have to replace the basic set $\mathbf{P}$ in the above definitions, with $\mathbf{P}_{p+q}$. Then each pair $(i, j) \in \mathbf{P}_{p+q}$ addresses a block instead of an element. Here, the pair (i, i) is disjoint with the pair (j, l) provided that $i \neq j$ and $i \neq l$.

2.4.4. Some classes of block-oriented pivot strategies

Here, we recall some special pivot strategies which have potential to be used in computation. They are efficient and make the standard symmetric Jacobi method provably convergent. Typically, on macro-level they use some known cyclic or quasi-cyclic strategy and on micro-level the row- or column-wise ordering. Later, we shall show that these *block-oriented pivot strategies* ensure the global convergence for the **J**-Jacobi methods. We list three important classes of block-oriented pivot strategies.

Class $\mathcal{A}_1$. Consider the cyclic block strategies defined by the *wave-front orderings* (see [30]) of $\mathbf{P}_{p+q}$. These are orderings of $\mathbf{P}_{p+q}$ from the relation (2.18) which are equivalent to $O_C(\mathbf{P}_{p+q})$. Let O be one such ordering. If we replace each pair (i, j), $i \leqslant j$ from O by either the sequence $O_C(A_{ij})$ or $O_R(A_{ij})$, we obtain an ordering of $\mathbf{P}$ which defines a cyclic pivot strategy from the class $\mathcal{A}_1$. We briefly say that the class $\mathcal{A}_1$ consists of those cyclic block-oriented pivot strategies, which are on macro-level equivalent to the column-cyclic strategy, and on micro-level they use the row- or column-wise ordering.

Class $\mathcal{A}_2$. These quasi-cyclic, block-oriented strategies have been recommended by Drmač and Veselić [14] as more efficient for computation than those from the class $\mathcal{A}_1$. They are defined in a similar way as the strategies from the class $\mathcal{A}_1$, except for the diagonal blocks A_{ii}, $1 \leqslant i \leqslant p + q$, which are operated twice within one quasi-sweep. A detailed consideration of these strategies is given in [20,21].

Class $\mathcal{A}_3$. These quasi-cyclic, block-oriented strategies are inspired by the quasi-cylic strategy proposed by Mascarenhas [26]. Let $2 \leqslant \alpha \leqslant n - 2$, $\beta = n - \alpha$, and let $\alpha_1, \alpha_2, \beta_1, \beta_2$ be non-negative integers such that $\alpha = \alpha_1 + \alpha_2$ and $\beta = \beta_1 + \beta_2$. Let us partition $\mathbf{P}$ into $\mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3, \mathbf{P}_4, \mathbf{P}_5, \mathbf{P}_6$ and $\mathbf{P}_7$, where

$$\begin{aligned}
\mathbf{P}_1 &= \{(i,j) : 1 \leqslant i < j \leqslant \alpha_1\},\\
\mathbf{P}_2 &= \{(i,j) : \alpha_1 + 1 \leqslant i < j \leqslant \alpha\},\\
\mathbf{P}_3 &= \{(i,j) : 1 \leqslant i \leqslant \alpha_1,\ \alpha_1 + 1 \leqslant j \leqslant \alpha\},\\
\mathbf{P}_4 &= \{(i,j) : \alpha + 1 \leqslant i < j \leqslant \alpha + \beta_1\},\\
\mathbf{P}_5 &= \{(i,j) : \alpha + \beta_1 + 1 \leqslant i < j \leqslant n\},\\
\mathbf{P}_6 &= \{(i,j) : \alpha + 1 \leqslant i \leqslant \alpha + \beta_1,\ \alpha + \beta_1 + 1 \leqslant j \leqslant n\},\\
\mathbf{P}_7 &= \{(i,j) : 1 \leqslant i \leqslant \alpha,\ \alpha + 1 \leqslant j \leqslant n\}.
\end{aligned}$$

Obviously, this partition must agree with the initial partition, so we assume $\alpha = \nu$ and $\beta = n - \nu$. Let

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{12}^* & \mathbf{A}_{22} \end{bmatrix} \begin{matrix} \} \alpha \\ \} \beta \end{matrix} = \begin{bmatrix} \overbrace{B_{11}}^{\alpha_1} & \overbrace{B_{12}}^{\alpha_2} & & \\ B_{12}^* & B_{22} & \mathbf{A}_{12} & \\ & & B_{33} & B_{34} \\ & \mathbf{A}_{12}^* & B_{34}^* & B_{44} \end{bmatrix} \begin{matrix} \} \alpha_1 \\ \} \alpha_2 \\ \} \beta_1 \\ \} \beta_2 \end{matrix} . \tag{2.19}$$

$$\underbrace{}_{\alpha} \quad \underbrace{}_{\beta} \qquad\qquad \underbrace{}_{\beta_1} \quad \underbrace{}_{\beta_2}$$

We see that the sets $\mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3, \mathbf{P}_4, \mathbf{P}_5, \mathbf{P}_6$, and $\mathbf{P}_7$ are associated with the submatrices $B_{11}, B_{22}, B_{12}, B_{33}, B_{44}, B_{34}$ and $\mathbf{A}_{12}$, respectively. Now, let O_i be $O_R(\mathbf{P}_i)$ or $O_C(\mathbf{P}_i)$ for $1 \leqslant i \leqslant 7$. Then the Mascarenhas quasi-cyclic strategy is defined by the ordering Θ, where

$$\Theta = [O_1, O_2, O_3, O_1, O_2, O_4, O_5, O_6, O_4, O_5, O_7].$$

Note that $\Theta \in \mathbf{O}(\mathbf{P})$ and all the elements of $\mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_4$ and $\mathbf{P}_5$ are repeated exactly twice in the sequence Θ. The sequence Θ defines a quasi-cylic pivot strategy which delivers, under some natural assumptions, the cubic asymptotic convergence per quasi-sweep (see [29]).
To obtain an associated quasi-cyclic, block-oriented strategy, all we need to assume and to do is the following:

(i) the basic partitions of $\mathbf{A}_{11}$ and $\mathbf{A}_{22}$ from (2.4) have to be sub-partitions of the partitions $\mathbf{A}_{11} = (B_{rs})$, $1 \leqslant r, s \leqslant 2$ and $\mathbf{A}_{22} = (B_{rs})$, $3 \leqslant r, s \leqslant 4$ from (2.19), respectively,
(ii) the serial ordering $O_R(A_{ij})$ or $O_C(A_{ij})$ has to be used for each block A_{ij}, $1 \leqslant i \leqslant j \leqslant p + q$, and
(iii) the block strategy (which selects the blocks A_{ij}), has to follow the ordering of pairs from Θ.

2.4.5. The corresponding classes of block pivot strategies

With each cyclic or quasi-cyclic block-oriented strategy is associated a *cyclic* or *quasi-cyclic block pivot strategy*. By this we simply mean the way of selecting the pivot blocks $A_{ij}^{(k)}$, $i = i(k), j = j(k)$. It is defined by an appropriate ordering or by a sequence of pairs from the set $\mathbf{P}_{p+q}$. Thus, with the class $\mathcal{A}_1$ is associated the class $\mathcal{B}_1$ which consists of (we can call them) wave-front block pivot strategies. With the classes $\mathcal{A}_2$ and $\mathcal{A}_3$ are associated the classes $\mathcal{B}_2$ and $\mathcal{B}_3$ of special quasi-cyclic block strategies. The classes $\mathcal{A}_1$–$\mathcal{A}_3$ ($\mathcal{B}_1$–$\mathcal{B}_3$) will be used together with *block steps* of block-oriented (*full block steps* of block) **J**-Jacobi methods. The full block steps and the corresponding methods are studied in [23].

3. The global convergence of block-oriented J-Jacobi methods

Here we prove the global convergence of the block-oriented **J**-Jacobi methods defined by (2.5) and by the pivot strategies from the classes $\mathcal{A}_1$–$\mathcal{A}_3$. To keep the exposition simple, we assume real matrices and real algorithms. This means $\Phi = I_n$ and in (2.10) as well as in the formulas for updating the diagonal elements (2.11) and (2.13), $|a_{rs}|$ has to be replaced by a_{rs}. All other relations from the previous subsection 2.3 hold. Since the strategies are from the classes $\mathcal{A}_1$–$\mathcal{A}_3$, each method is cyclic or quasi-cyclic with period M. We use N for $N_n = n(n-1)/2$.

The following observation will simplify the analysis.

Lemma 3.1. *For each strategy $\mathcal{I}$ from the classes $\mathcal{A}_1$ or $\mathcal{A}_3$, there exists an equivalent strategy $\mathcal{I}_0$, such that within one sweep or quasi-sweep all the hyperbolic transformations are performed consecutively, using the column-wise ordering $O_C(\mathbf{A}_{12})$. In another words, $O_{\mathcal{I}} \sim O_{\mathcal{I}_0} = [O'_{\mathcal{I}_0}, O_C(\mathbf{A}_{12}), O''_{\mathcal{I}_0}]$ where $O'_{\mathcal{I}_0}$ or $O''_{\mathcal{I}_0}$, but not both of them, can be empty.*

Proof. We shall use the following notation: $\mathcal{C}_{ij} = \mathcal{O}_C(A_{ij})$ for all $1 \leqslant i \leqslant j \leqslant p+q$.

If $\mathcal{I} \in \mathcal{A}_3$, the claim is obvious for four reasons:

(1) The last $\nu(n-\nu)$ pairs from $O_{\mathcal{I}}$ and only they address the elements of $\mathbf{A}_{12}$,
(2) $O_R(A_{ij}) \sim O_C(A_{ij}) = \mathcal{C}_{ij}$ for all $i \leqslant j$,
(3) $[\mathcal{C}_{1,p+1}, \mathcal{C}_{1,p+2}, \ldots, \mathcal{C}_{1,p+q}, \mathcal{C}_{2,p+1}, \ldots, \mathcal{C}_{2,p+q}, \mathcal{C}_{p,p+1}, \ldots, \mathcal{C}_{p,p+q}] \sim$
$[\mathcal{C}_{1,p+1}, \mathcal{C}_{2,p+1}, \ldots, \mathcal{C}_{p,p+1}, \mathcal{C}_{1,p+2}, \ldots, \mathcal{C}_{p,p+2}, \mathcal{C}_{1,p+q}, \ldots, \mathcal{C}_{p,p+q}]$,
(4) $[\mathcal{C}_{1,p+1}, \mathcal{C}_{2,p+1}, \ldots, \mathcal{C}_{p,p+1}, \mathcal{C}_{1,p+2}, \ldots, \mathcal{C}_{p,p+2}, \mathcal{C}_{1,p+q}, \ldots, \mathcal{C}_{p,p+q}] \sim O_C(\mathbf{A}_{12})$.

Let $\mathcal{I} \in \mathcal{A}_1$. Then it is equivalent to the cyclic strategy $\mathcal{I}_1$, which is on macro-level column-cyclic and on micro-level (i.e., within each block) uses column-wise ordering. It is easy to see that $\mathcal{I}_1 \sim \mathcal{I}_C$ where $\mathcal{I}_C$ is the column-cyclic strategy. However, the column-cyclic strategy is equivalent to $\mathcal{I}_0$ which is defined by the sequence of pairs $[O_C(\mathbf{A}_{11}), O_C(\mathbf{A}_{12}), O_C(\mathbf{A}_{22})]$. Since $\sim$ is equivalence relation, we have $\mathcal{I} \sim \mathcal{I}_0$. □

From the proof of Lemma 3.1 we see that each strategy from the class $\mathcal{A}_1$ is equivalent to the column-cyclic strategy. Now, using [33, Theorem 2.3] and Theorem 3.7, we obtain convergence for the class $\mathcal{A}_1$. However, for the completeness of exposition, we shall use another, unified approach which will deliver convergence for all three classes $\mathcal{A}_1$–$\mathcal{A}_3$.

The next lemma actually means: $\lim_{\tau\to\infty} \mathbf{A}_{12}^{(\tau M)} = 0$, i.e., the $(1,2)$-block of the iterate $\mathbf{A}^{(\tau M)}$, which is obtained at the end of the τth sweep or quasi-sweep, tends to zero as τ increases.

Lemma 3.2. *For any pivot strategy from the classes $\mathcal{A}_1$–$\mathcal{A}_3$ and for any $\varepsilon > 0$ there exists $\tau_0 \geqslant 0$, such that*

$$\|\mathbf{A}_{12}^{(\tau M)}\| \leqslant \varepsilon, \quad \tau \geqslant \tau_0. \tag{3.1}$$

Proof. The proof is lengthy and technical, so it is moved to Appendix. □

In order to prove appropriate results for $(1,1)$- and $(2,2)$-blocks, we shall need the notion of *ζ-convergent* periodic strategies.

Definition 3.3. Let $\mathcal{I}$ be a periodic strategy with period M and let $\zeta \geqslant 0$ be a real number. Let θ_k, $0 \leqslant k \leqslant M-1$ be arbitrary M real numbers satisfying the condition $|\theta_k| \leqslant \zeta$, $0 \leqslant k \leqslant M-1$. Let $H = H^{(0)}$ be any symmetric matrix of order n. Consider the iterative process

$$H^{(k+1)} = \widetilde{H}^{(k)} - \tilde{h}^{(k)}_{i(k)j(k)} e_{i(k)} e_{j(k)}^T - \tilde{h}^{(k)}_{j(k)i(k)} e_{j(k)} e_{i(k)}^T, \quad k = 0, 1, \ldots, M-1,$$

where

$$\widetilde{H}^{(k)} = R^T_{i(k)j(k)}(\theta_k) H^{(k)} R_{i(k)j(k)}(\theta_k) = (\tilde{h}^{(k)}_{st}). \tag{3.2}$$

Here $R_{i(k)j(k)}(\theta_k)$ is the rotation in the $(i(k), j(k))$-plane with angle θ_k, and $(i(k), j(k)) = \mathcal{I}(k)$ for all $0 \leqslant k \leqslant M-1$. The strategy $\mathcal{I}$ is *ζ-convergent* if

$$\mathrm{Off}(H^{(M)}) \leqslant \mu \mathrm{Off}(H), \quad 0 \leqslant \mu < 1,$$

where μ depends only on n and ζ.

Henrici and Zimmermann [24] have shown for the column-cyclic strategy

$$\mathrm{Off}^2(H^{(N)}) \leqslant \left(1 - \prod_{\substack{1 \leqslant i < j \leqslant n \\ i < j-1}} \cos^2(\theta_{ij})\right) \mathrm{Off}^2(H),$$

where θ_{ij} is the rotation angle used by R_{ij} which rotates in the (i,j)-plane. This implies that one can take

$$\mu = \sqrt{1 - (\cos\zeta)^{(n^2-3n+2)/2}}.$$

Thus, the column-cyclic strategy and therefore any strategy from the class $\mathcal{A}_1$ is ζ-convergent if $\zeta < \pi/2$.

By inspecting the proof of [20, Theorem 2.1], one can verify that for any basic partition, $\mathcal{I}_{\mathcal{A}_2}$ (see the proof of Lemma 3.2) is ζ-convergent if $\zeta \leqslant \pi/4$. This implies that any strategy from the class $\mathcal{A}_2$ is ζ-convergent with $\zeta \leqslant \pi/4$. Inspecting the proof means to check whether just the condition $|\theta_k| \leqslant \zeta, 0 \leqslant k \leqslant M-1$, plus the annihilation of the pivot elements is used.

One of our main tools for proving convergence is the following result.

Lemma 3.4. *Let H and $F^{(k)}, k \geqslant 0$ be symmetric matrices of order n and let $\mathcal{I}$ be a periodic strategy with period M. Let the sequence $H^{(0)} = H, H^{(1)}, \ldots$ be obtained by the iterative process*

$$H^{(k+1)} = [R_{i(k)j(k)}(\theta_k)]^T H^{(k)} R_{i(k)j(k)}(\theta_k) + F^{(k)}, \quad k \geqslant 0, \tag{3.3}$$

where $R_{i(k)j(k)}(\theta_k)$ is a rotation in the $(i(k), j(k))$-plane where $(i(k), j(k)) = \mathcal{I}(k), k \geqslant 0$. Let the following conditions hold

(i) $\tilde{h}^{(k)}_{i(k)j(k)} \to 0$ *as* $k \to \infty$, *where* $\widetilde{H}^{(k)}$ *is as in* (3.2),
(ii) $\mathcal{I}$ *is* ζ*-convergent and*
(iii) $|\theta_k| \leqslant \zeta, k \geqslant k_0$ *for some* $k_0 \geqslant 0$.

Then the following two conditions are equivalent

(a) $\mathrm{Off}(F^{(k)}) \to 0$ *as* $k \to \infty$,
(b) $\mathrm{Off}(H^{(k)}) \to 0$ *as* $k \to \infty$.

Proof. The proof is a slight generalization of the proof of [17, 2.2 Lemma]. We only have to show that the spectral norm of the Jacobi operator $\Im_\tau$ defined by $\mathcal{I}$ and by the finite sequence of rotations $(R_{i(k)j(k)}(\theta_k), (\tau-1)M \leqslant k < M)$, is bounded (uniformly, with respect to $\tau, \tau \geqslant \tau_0 \geqslant k_0/M$) by some constant $\mu, 0 \leqslant \mu < 1$, which depends on n and ζ. However, this is implied by the assumptions (ii) and (iii). One just has to consider the subsequence $(H^{(k)}, k \geqslant k_0)$. □

The next two lemmas prove that the off-norms of $\mathbf{A}_{11}^{(\tau M)}$ and $\mathbf{A}_{22}^{(\tau M)}$ tend to zero as τ increases. Their proofs can be found in Appendix. The main ideas in the proofs are as follows. For each class, $\mathcal{A}_1$, $\mathcal{A}_2$ or $\mathcal{A}_3$, find a suitable strategy, such that under that strategy, the process on $\mathbf{A}^{(k)}$ can be "restricted" on $\mathbf{A}_{11}^{(k)}$ and $\mathbf{A}_{22}^{(k)}$. And thanks to Lemma 3.1, the induced processes can be viewed in the form (3.3), with the conditions (i)–(iii) and (a) of Lemma 3.4 satisfied.

Lemma 3.5. *For any pivot strategy from the classes $\mathcal{A}_1$–$\mathcal{A}_3$ and for any $\varepsilon > 0$, there exists $\tau_0 \geqslant 0$, such that*

$$\mathrm{Off}(\mathbf{A}_{11}^{(\tau M)}) \leqslant \varepsilon, \qquad \tau \geqslant \tau_0.$$

Lemma 3.6. *For any pivot strategy from the classes $\mathcal{A}_1$–$\mathcal{A}_3$ and for any $\varepsilon > 0$, there exists $\tau_0 \geqslant 0$, such that*

$$\mathrm{Off}(\mathbf{A}_{22}^{(\tau M)}) \leqslant \varepsilon, \qquad \tau \geqslant \tau_0.$$

Theorem 3.7. *For any pivot strategy from the classes $\mathcal{A}_1$–$\mathcal{A}_3$,*

$$\mathbf{A}^{(k)} \to \mathbf{\Lambda} \quad \text{as} \quad k \to \infty.$$

Proof. By Lemmas 3.2, 3.5, and 3.6, we know that $\mathrm{Off}(\mathbf{A}^{(\tau M)}) \to 0$ as $\tau \to \infty$. For large enough τ, the effect of the hyperbolic rotations on $\mathrm{Off}(\mathbf{A}^{((\tau-1)M)})$ is negligible, and orthogonal rotations do not increase it. Hence, there exist non-negative real numbers ν_τ, $\tau \geqslant 1$, such that

$$\mathrm{Off}(\mathbf{A}^{(k)}) \leqslant \mathrm{Off}(\mathbf{A}^{((\tau-1)M)}) + \nu_\tau, \quad (\tau - 1)M \leqslant k \leqslant \tau M,$$

and $\nu_\tau \to 0$ as $\tau \to \infty$. This shows that $\mathrm{Off}(\mathbf{A}^{(k)}) \to 0$ as $k \to \infty$.

It remains to show that $\mathrm{diag}(\mathbf{A}^{(k)}) \to \mathbf{\Lambda}$ as $k \to \infty$.

We shall actually show that $\mathbf{J}\,\mathrm{diag}(\mathbf{A}^{(k)}) \to \mathbf{J\Lambda}$ as $k \to \infty$. Perhaps the simplest way to prove it is by using the theory of Gerschgorin discs. Let

$$\begin{aligned}\lambda_1 = \lambda_2 = \cdots = \lambda_{s_1} > \lambda_{s_1+1} = \cdots = \lambda_{s_2} > \cdots > \lambda_{s_{u-1}+1} = \cdots = \lambda_{s_u} \\ > -\lambda_{s_u+1} = \cdots = -\lambda_{s_{u+1}} > \cdots > -\lambda_{s_{\omega-1}+1} = \cdots = -\lambda_{s_\omega}\end{aligned}$$

be the eigenvalues of **JA**. Then $\mathtt{n}_i = s_i - s_{i-1}$, $1 \leqslant i \leqslant \omega$ are their multiplicities. Let

$$\delta = \min\left\{ \min_{1 \leqslant i \leqslant u-1} (\lambda_{s_i} - \lambda_{s_{i+1}}), \lambda_{s_u} + \lambda_{s_{u+1}}, \min_{u+1 \leqslant i \leqslant \omega-1} (\lambda_{s_{i+1}} - \lambda_{s_i}) \right\}$$

be the minimum gap between two different eigenvalues of **JA**. Consider the discs $\mathcal{D}_i$, $1 \leqslant i \leqslant \omega$ of radius $\gamma = \delta/3$ around the eigenvalues of **JA**. Since $\lim_{k\to\infty} \mathrm{Off}(\mathbf{A}^{(k)}) = 0$, the Gerschgorin theory implies that there is k_0 such that for every k, $k \geqslant k_0$, each disc $\mathcal{D}_i$ contains exactly $\mathtt{n}_i$ diagonal elements of $\mathbf{JA}^{(k)}$, and we can additionally require that the pivot elements satisfy $|a_{ij}^{(k)}| \leqslant \gamma/2$.

Indeed, for this conclusion it suffices to assume that k_0 is so large that all Gerschgorin discs of $\mathbf{JA}^{(k)}$ have radii not larger than $\gamma/2n$. Then $|a_{rs}^{(k)}| \leqslant \gamma/2$, $r \neq s$ and each eigenvalue λ_{s_i} or $-\lambda_{s_i}$ of $\mathbf{JA}^{(k)}$ must lie in the connected component which consits of precisely $\mathtt{n}_i$ Gerschgorin discs. Any other possibility would contradict the Gerschgorin theory.

Namely, each connected component of Gerschgorin discs contains as many eigenvalues (counting multiplicities) as many discs it contains. For given i, at least one Gerschgorin disc contains λ_{si} and therefore the whole connected component to which it belongs must be contained within $\mathcal{D}_i$. The same holds for $-\lambda_{si}$.

To end the proof, it suffices to show that no diagonal element can change its affiliation with one of the discs $\mathcal{D}_i$, $1 \leqslant i \leqslant \omega$ during the kth step. This is sufficient since as k increases, the radii of Gerschgorin discs shrink to zero, so all the diagonals of $\mathbf{JA}^{(k)}$ contained within $\mathcal{D}_i$ must tend to its center which is the eigenvalue of **JA**.

If the kth rotation is hyperbolic, then the affected diagonals $a_{ii}^{(k)}$ and $a_{jj}^{(k)}$ must lie in different discs, say $\mathcal{D}_\alpha$ and $\mathcal{D}_\beta$, respectively. But the maximum shift allowed for the change of the diagonal elements is bounded by $|\tanh\theta| \cdot |a_{ij}^{(k)}| \leqslant 1 \cdot \gamma/2$. So, they cannot interchange the discs $\mathcal{D}_\alpha$ and $\mathcal{D}_\beta$, respectively.

Let the kth rotation be orthogonal, and let the changing diagonals $a_{ii}^{(k)}$ and $a_{jj}^{(k)}$ lie in the discs $\mathcal{D}_\alpha$ and $\mathcal{D}_\beta$. If $\alpha \neq \beta$, then the argument is similar as above, since the maximum allowed shifts are $\pm\tan\theta\, a_{ij}^{(k)}$ and $|\tan\theta| \leqslant 1$. If $\alpha = \beta$, then the both diagonals remain in the same disc $\mathcal{D}_\alpha$. Thus in all cases the affiliation is preserved. □

3.1. Asymptotic convergence

Here we address the way how $\mathrm{Off}(\mathbf{A}^{(rM)})$ behaves when $\mathrm{Off}(\mathbf{A}^{((r-1)M)})$ becomes small enough (say, smaller than δ, the minimum gap in the spectrum of **A**). The scope of this paper allows us just to say what is known and what can be expected to hold.

In order to ensure the quadratic convergence for the case of multiple eigenvalues, one has to ensure (cf. [21]) that the basic partition $(n_1, \ldots, n_{p+q})$ and the *natural partition* $(\mathtt{n}_1, \ldots, \mathtt{n}_\omega)$ (see the proof of Theorem 3.7) are sub-partitions of the initial partition. For the natural partition, which is determined by the multiplicities of the eigenvalues of **JA**, this condition is automatically fulfilled, so care must be

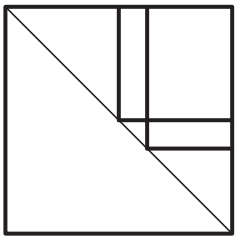

Fig. 4.1. Three typical loops in the two-sided Jacobi algorithm.

taken for the basic partition. The general proof seems to be difficult to make, because of the interplay between the basic and the natural partition. In other words, for each special relation between these two partitions, a separate proof may be needed.

The asymptotic convergence of the **J**-Jacobi methods defined by strategies from the class $\mathcal{A}_1$ is well understood. The quadratic convergence of ordinary iterates per sweep has been proved in [11] and of scaled iterates in [27].

The asymptotic quadratic convergence per quasi-sweep of the symmetric Jacobi methods defined by strategies from the class $\mathcal{A}_2$ has been considered in [21]. It seems that quadratic convergence is always present. The same should hold for the **J**-Jacobi methods considered here.

Finally, in [29] the cubic convergence per quasi-sweep has been proved for the symmetric Jacobi method defined by a strategy from the class $\mathcal{A}_3$ (actually, for the case $n_i = 1, 1 \leqslant i \leqslant n$). We are confident, that a similar proof can be made for the **J**-Jacobi methods considered here.

4. Computational approach to the one-sided block J- Jacobi algorithm

The block-Jacobi methods are usually described in the context of parallel computing, see for example [28,1]. Here, instead of analyzing how to decouple the processes, we consider how to reuse data when they reach the cache memory. This consideration can be used in the construction of a parallel Jacobi method.

After the Hermitian indefinite factorization (1.1) is completed, the one-sided Jacobi algorithm is applied to compute the HSVD of the factor G. The block partition (2.4) induces the block column-partition of G and the corresponding partition of the diagonal matrix **J**,

$$G = [G_1, \ldots, G_p, G_{p+1}, \ldots, G_{p+q}], \qquad \mathbf{J} = \operatorname{diag}(I_{n_1}, \ldots, I_{n_p}, -I_{n_{p+1}}, \ldots, -I_{n_{p+q}}),$$

where G_i has n_i columns for $1 \leqslant i \leqslant p+q$.

At step k, the one-sided Jacobi method computes the $\widehat{\mathbf{J}}$-unitary matrix $\widehat{\mathbf{V}}^{(k)}$ (see the relations (2.5)–(2.7)) to make the block-columns G_i and G_j more orthogonal. After the completion of the orthogonalization process, all columns of G have to be numerically orthogonal, i.e., $G \cdot (\prod_k \mathbf{V}^{(k)}) \approx U\Sigma$ has to hold. Having the computed $U\Sigma$ at hand, we can easily extract the squares of the hyperbolic singular values σ_i^2 and form $\sigma_i^2 j_{ii}$, which are the eigenvalues of H.

4.1. The block-oriented algorithm

In a block-oriented algorithm, all off-diagonal elements of **A** are annihilated once or twice during one sweep or quasi-sweep. The quantities computed at step k of the algorithm will be indexed by the superscript (k).

Let a cyclic pivot strategy be used. Then all elements of each off-diagonal block are annihilated once. At step k we form the matrix

$$\widehat{\mathbf{A}}^{(k)} = \begin{bmatrix} A_{ii}^{(k)} & A_{ij}^{(k)} \\ A_{ji}^{(k)} & A_{jj}^{(k)} \end{bmatrix} = \begin{bmatrix} \left[G_i^{(k)}\right]^* G_i^{(k)} & \left[G_i^{(k)}\right]^* G_j^{(k)} \\ \left[G_j^{(k)}\right]^* G_i^{(k)} & \left[G_j^{(k)}\right]^* G_j^{(k)} \end{bmatrix}.$$

Then $n_i \cdot n_j$ two-sided (unitary or hyperbolic) Jacobi steps are applied to $\widehat{\mathbf{A}}^{(k)}$ to sequentially annihilate the elements of $A_{ij}^{(k)}$. The transformations are accumulated in the matrix $\widehat{\mathbf{V}}^{(k)}$ and then $[G_i^{(k)}, G_j^{(k)}]$ is post-multiplied by $\widehat{\mathbf{V}}^{(k)}$.

However, this way is slow in computational time. Direct diagonalization of $\widehat{A}^{(k)}$ involves essentially 3 loops, and only one of them is fast (the one starting with indices $(i, 1)$ and $(j, 1)$, see Fig. 4.1). A better approach (nearly 2.5 times faster even for small matrices) is to factor $\widehat{\mathbf{A}}^{(k)}$ by the Cholesky factorization, $\widehat{\mathbf{A}}^{(k)} = [R^{(k)}]^* R^{(k)}$, and apply $n_i \cdot n_j$ one-sided Jacobi steps on $R^{(k)}$.

We note that we could have obtained the same factor $R^{(k)}$ by using the QR factorization of the matrix $[G_i^{(k)}, G_j^{(k)}]$, but this way, although more accurate, is too slow. For our algorithm it is important that the used transformation matrix $\widehat{\mathbf{V}}^{(k)}$ is as close as possible to a $\widehat{\mathbf{J}}$-unitary matrix. The matrix $\widehat{\mathbf{V}}^{(k)}$ makes the columns of $R^{(k)}$, and thus of $[G_i^{(k)}, G_j^{(k)}]$ more orthogonal.

The next problem is how to compute the transformation

$$[G_i^{(k+1)}, G_j^{(k+1)}] = [G_i^{(k)}, G_j^{(k)}]\widehat{\mathbf{V}}^{(k)}$$

using the BLAS 3 routines. Note that the BLAS 3 multiplication routine xGEMM cannot compute $[G_i^{(k+1)}, G_j^{(k+1)}]$ in place of $[G_i^{(k)}, G_j^{(k)}]$. So, we need extra workspace $[X, Y]$ of dimension $n \times 2n_{\max}$ where $n_{\max} = \max_{1 \leqslant r \leqslant p+q}\{n_r\}$. Typically, we need four xGEMM operations:

$$\begin{aligned} X &= G_j^{(k)} V_{21}^{(k)}, & Y &= G_j^{(k)} V_{22}^{(k)}, \\ X &= G_i^{(k)} V_{11}^{(k)} + X, & Y &= G_i^{(k)} V_{12}^{(k)} + Y, \end{aligned}$$

so that $G_i^{(k+1)}$ ($G_j^{(k+1)}$) is stored in X (Y). Now a new workspace (playing the role of X, Y) is freed at the position previously occupied by $G_i^{(k)}$ and $G_j^{(k)}$. To minimize the number of copies needed, we keep track of the positions where each $G_i^{(k+1)}$ ended. At the final stage of the whole process the block-columns are re-permuted to their proper places.

A similar approach is used for the diagonal blocks of $\mathbf{A}^{(k)}$. Here, one computes the Cholesky factor $R^{(k)}$ of

$$\widehat{\mathbf{A}}^{(k)} = A_{ii}^{(k)} = [G_i^{(k)}]^* G_i^{(k)}$$

and applies to it a full sweep of the serial one-sided Jacobi method.

Finally, let us briefly comment the relative accuracy result from [23] which also applies to this one-sided algorithm. That result states that the maximum relative error in the computed singular values, coming from one block step, is bouned by $\gamma_k \cdot \kappa_2(C^{(k)}) \cdot \varepsilon$. Here ε is the machine round-off,

$$G^{(k)}\mathbf{V}^{(k)} = \Delta^{(k)} C^{(k)}$$

and

$$\gamma_k = f(n)\kappa_2(\mathbf{V}^{(k)})\kappa_2(V^{(k)}),$$

where f is a slowly growing function of n, $\kappa_2(\mathbf{V}^{(k)})$ is the spectral condition of the transformation matrix, and $\kappa_2(V^{(k)})$ is the condition of the J-unitary matrix from the hyperbolic SVD of the current iterate $G^{(k)}$. The diagonal matrix $\Delta^{(k)}$ can be chosen to minimize $\kappa_2(C^{(k)})$. Therefore, $\kappa_2(C^{(k)})$ can be arbitrarily smaller than $\kappa_2(G^{(k)}\mathbf{V}^{(k)})$, which makes the estimate attractive. As the process advances, $\kappa_2(V^{(k)})$ and $\kappa_2(\mathbf{V}^{(k)})$ tend to one.

5. Numerical examples

We have made our tests on Pentium 4 computer running under Windows XP Professional x64 edition. We have used Intel FORTRAN compiler version 9.1.028, and BLAS and LAPACK routines contained in Intel Math Kernel Library 8.1. Our compiler optimization was set to aggressive optimization (`fast`,

Table 5.1
Compiler optimized simple algorithm versus the corresponding simple BLAS 1 algorithm.

$n = 1000$								
Block size n_i	8	16	24	32	40	48	56	64
Gain (%)	13.38	14.12	17.82	15.89	13.95	12.67	8.63	11.40
Block size n_i	72	80	88	96	104	112	120	128
Gain (%)	9.69	7.68	7.44	10.35	6.80	6.18	2.87	5.80

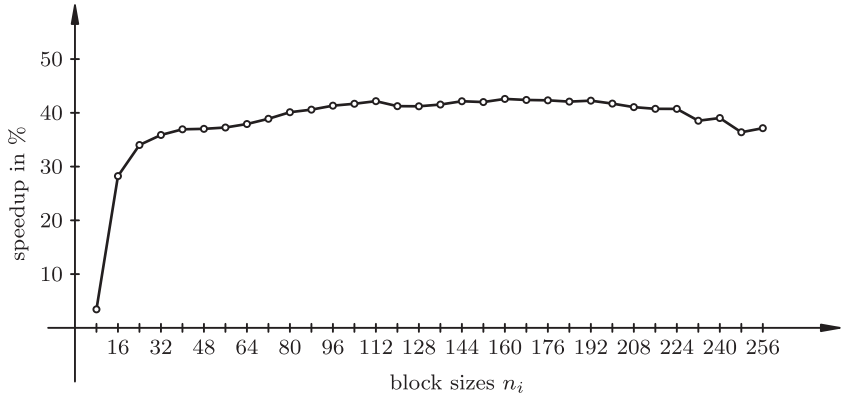

Fig. 5.1. Speedup of the column-cyclic block-oriented J-Jacobi method over the non-blocked counterpart, matrix size $n = 4000$.

`optimize:5`). Under these circumstances, our "small", in cache Jacobi algorithm (operating on $\widehat{\mathbf{A}}^{(k)}$) performs somewhat better if the program segment is not written using the BLAS routines. This is so because the compiler vectorizes the loops and the gain is better than using the BLAS (see Table 5.1).

We have tested the block algorithms on real matrices in double precision. The elements of the upper-triangle of test symmetric matrices have been randomly generated, using the LAPACK routine `DLARND` (uniform distribution with elements in $[-5, 5]$). We have tested matrices of order from 500 to 4000 in steps of 500.

Our tests show that the non-blocked Jacobi algorithm runs faster than its block counterparts for matrices of order less than 1000. For these matrices a big portion of them resides in the cache memory, while book-keeping of smaller matrices ($\widehat{\mathbf{A}}$, $\widehat{\mathbf{V}}$, $[G_i\ G_j]$ ...) which are part of the block algorithm takes time. We have tested algorithms with equally-sized block-columns G_i (except for the last block-column) each consisting of 8–256 columns.

The tests show that the reduction in computational time for the block-oriented algorithm increases up to 40% for the block dimensions of order 32–256, provided that initially the symmetric indefinite factorization with complete pivoting is used. It is interesting that the block dimension almost does not influence acceleration of the algorithm. If we replace the complete pivoting with the partial pivoting, the whole diagonalization process is a bit slower. The Fig. 5.1 displays the speedups of the block oriented algorithm compared to the non-blocked algorithm for $n = 4000$.

For matrices of order between 1500 and 3500, the figures (like Fig. 5.1) have the same shape, but the speedups are lower.

Our experiments have shown that in the case of an indefinite matrix H, the number of positive eigenvalues does not influence the performance. However, if H is definite, the method is faster. For the matrices of order 1500, the method is approximately 25% slower for indefinite matrices than for the definite ones.

In all our experiments, the accuracy of the hyperbolic singular values computed by the block-oriented methods, was approximately the same as for those computed by their non-blocked counterparts.

6. Conclusion and future work

In this paper we have considered a way how to accelerate the accurate eigensolver of Veselić [33] for indefinite Hermitian matrices, by transforming it into a BLAS 3 algorithm. For the block-oriented algorithms we have proved the global convergence and briefly discussed the asymptotic convergence. It is known that the algorithms are relatively accurate [23]. Numerical tests show that the block-oriented algorithms can be, for larger matrices, more than 40% faster than the non-blocked algorithms.

All these results should encourage further research in several directions. There are several possibilities how to further accelerate the block-oriented algorithm. In the iterative part of the algorithm, one can use ideas from [19,18] to apply fast scaled block transformations. In the preprocessing part one can try to devise a BLAS 3 modification of the indefinite QR factorization [31] if the factor G is given. Finally, one can try to adapt some of the many tricks advocated in [13,14]. If accuracy is the predominant issue, it is important to devise a criterion, perhaps after computing G, whether to apply our block method to G or to return to H and proceed with the orthogonal method from [10].

Acknowledgement

The authors are thankful to the anonymous referees for the valuable comments which improved the paper.

Appendix A

Here, in the proofs of Lemmas 3.2, 3.5, and 3.6, we use notation $\mathbf{R}^{(k)}$ instead of $V^{(k)}$ for the hyperbolic or trigonometric plane rotation at step k. This is in accordance with the notation $\widehat{V} = \widehat{\Phi}\widehat{R}$ (see also the relation (2.10)) which has been introduced in Section 2.3. Since we deal with real matrices and processes, we have at each step $\widehat{\Phi} = I_n$.

Note that R ($R^{(k)}$) has been used in Introduction and in Section 2.1 (in Section 4.1) for the triangular QR factor of G (of $[G_i^{(k)}, G_j^{(k)}]$).

Proof of Lemma 3.2. Let us first assume that $\mathcal{I}$ belongs to the class $\mathcal{A}_1$ or $\mathcal{A}_3$. By the previous lemma, there is an equivalent strategy $\mathcal{I}_0$ with the special property. Since the both strategies $\mathcal{I}$ and $\mathcal{I}_0$ yield the same matrix $\mathbf{A}^{(\tau M)}$ at the end of cycle τ, we can assume that the process is defined by $\mathcal{I}_0$.

Let us consider the τth sweep or quasi-sweep. First note that orthogonal transformations whose pivot elements lie within the blocks $\mathbf{A}_{11}^{(k)}$ or $\mathbf{A}_{22}^{(k)}$ do not change $\|\mathbf{A}_{12}^{(k)}\|$. So, we consider only the hyperbolic steps k such that $(\tau-1)M+1 \leqslant k \leqslant \tau M$ and we consider what happens with $\|\mathbf{A}_{12}^{(k)}\|$ when τ grows.

For such a k, the hyperbolic transformation has the form $\mathbf{A}^{(k+1)} = [\mathbf{R}^{(k)}]^T\mathbf{A}^{(k)}\mathbf{R}^{(k)}$. Since the pivot elements $a_{rs}^{(k)}$ and $a_{sr}^{(k)}$ ($= a_{rs}^{(k)}$) are annihilated, we can write

$$\mathbf{A}^{(k+1)} = [\mathbf{R}^{(k)}]^T\mathbf{A}^{(k)}\mathbf{R}^{(k)} = \mathbf{A}^{(k)} - a_{rs}^{(k)}e_re_s^T - a_{rs}^{(k)}e_se_r^T + F^{(k)}, \tag{A.1}$$

where e_i is the ith column of I_n. In (A.1) the (r,s)- and (s,r)-elements of $F^{(k)}$ are zero. Because of (2.17), $\mathbf{R}^{(k)} \to I_n$ and $a_{rs}^{(k)} \to 0$ as $\tau \to \infty$. Hence, as $\tau \to \infty$, the product of all hyperbolic rotations appearing in the sweep or quasi-sweep τ, tends to the identity matrix. By (2.14), $\|\mathbf{A}^{(k)}\| \leqslant \mathrm{tr}(A)$ holds for all k, so we conclude that $\|F^{(k)}\|$ can be made arbitrarily small for large enough τ.

Let us iterate the relation (A.1) over all hyperbolic steps k within the sweep or quasi-sweep τ. Because $\mathcal{I}_0$ has the special property described in Lemma 3.1, there exist positive integers k'_τ and k''_τ such that $(\tau-1)M+1 \leqslant k'_\tau \leqslant k \leqslant k''_\tau = k'_\tau + \nu(n-\nu) \leqslant \tau M$. Since the elements of $\mathbf{A}_{12}^{(k'_\tau)}$ are annihilated in the column-wise fashion, after applying the first two hyperbolic transformations, we have

$$\begin{aligned}\mathbf{A}^{(k'_\tau+2)} &= \mathbf{A}^{(k'_\tau+1)} - a_{2,m+1}^{(k'_\tau+1)}e_2e_{m+1}^T - a_{m+1,2}^{(k'_\tau+1)}e_{m+1}e_2^T + F^{(k'_\tau+1)}\\ &= \mathbf{A}^{(k'_\tau)} - a_{1,m+1}^{(k'_\tau)}e_1e_{m+1}^T - a_{m+1,1}^{(k'_\tau)}e_{m+1}e_1^T + F^{(k'_\tau)}\end{aligned}$$

$$
\begin{aligned}
&- a_{2,m+1}^{(k'_\tau+1)} e_2 e_{m+1}^T - a_{m+1,2}^{(k'_\tau+1)} e_{m+1} e_2^T + F^{(k'_\tau+1)} \\
&= \mathbf{A}^{(k'_\tau)} - a_{1,m+1}^{(k'_\tau)} e_1 e_{m+1}^T - a_{m+1,1}^{(k'_\tau)} e_{m+1} e_1^T \\
&\quad - a_{2,m+1}^{(k'_\tau)} e_2 e_{m+1}^T - a_{m+1,2}^{(k'_\tau)} e_{m+1} e_2^T + \mathbf{F}^{(k'_\tau+2)}
\end{aligned}
$$

where $\|\mathbf{F}^{(k'_\tau+2)}\|$ can be made arbitrarily small for large enough τ. Continuing this process, after $\nu(n-\nu)$ hyperbolic steps, we obtain

$$
\mathbf{A}^{(k''_\tau)} = \begin{bmatrix} \mathbf{A}_{11}^{(k'_\tau)} & 0 \\ 0 & \mathbf{A}_{22}^{(k'_\tau)} \end{bmatrix} + \mathbf{F}^{(k''_\tau)} = \begin{bmatrix} \mathbf{A}_{11}^{(k''_\tau)} & \mathbf{F}_{12}^{(k''_\tau)} \\ [\mathbf{F}_{12}^{(k''_\tau)}]^T & \mathbf{A}_{22}^{(k''_\tau)} \end{bmatrix} \tag{A.2}
$$

and for sufficiently large τ, $\|\mathbf{F}^{(k''_\tau)}\|$ can be made arbitrarily small. The following $\tau M - k''_\tau$ orthogonal steps will not change $\|\mathbf{A}_{12}^{(k''_\tau)}\|$. Therefore, $\|\mathbf{A}_{12}^{(\tau M)}\| = \|\mathbf{F}_{12}^{(k''_\tau)}\|$, implying the assertion (3.1).

It remains to address the class of quasi-cyclic strategies $\mathcal{A}_2$. Each strategy $\mathcal{I}$ from this class is equivalent to the strategy $\mathcal{I}_{\mathcal{A}_2}$, which is defined as follows (see [20]):

$$
\begin{aligned}
\mathcal{I}_{\mathcal{A}_2} = {} & \mathcal{C}_{11}, \mathcal{C}_{11}, \mathcal{C}_{22}, \mathcal{C}_{12}, \mathcal{C}_{22}, \mathcal{C}_{13}, \mathcal{C}_{33}, \mathcal{C}_{23}, \mathcal{C}_{33}, \ldots, \\
& \mathcal{C}_{1p}, \mathcal{C}_{2p}, \ldots, \mathcal{C}_{p-2,p}, \mathcal{C}_{pp}, \mathcal{C}_{p-1,p}, \mathcal{C}_{pp}, \\
& \mathcal{C}_{1,p+1}, \mathcal{C}_{2,p+1}, \ldots, \mathcal{C}_{p-1,p+1}, \mathcal{C}_{p+1,p+1}, \mathcal{C}_{p,p+1}, \mathcal{C}_{p+1,p+1}, \ldots \\
& \mathcal{C}_{1,p+q}, \mathcal{C}_{2,p+q}, \ldots, \mathcal{C}_{p+q-2,p+q}, \mathcal{C}_{p+q,p+q}, \mathcal{C}_{p+q-1,p+q}, \mathcal{C}_{p+q,p+q}.
\end{aligned}
$$

It is easy to check that $\mathcal{I}_{\mathcal{A}_2}$ is equivalent to

$$
\begin{aligned}
\mathcal{I}'_{\mathcal{A}_2} = {} & \mathcal{C}_{11}, \mathcal{C}_{11}, \mathcal{C}_{22}, \mathcal{C}_{12}, \mathcal{C}_{22}, \mathcal{C}_{13}, \mathcal{C}_{33}, \mathcal{C}_{23}, \mathcal{C}_{33}, \ldots, \\
& \mathcal{C}_{1p}, \mathcal{C}_{2p}, \ldots, \mathcal{C}_{p-2,p}, \mathcal{C}_{pp}, \mathcal{C}_{p-1,p}, \mathcal{C}_{pp}, \\
& \mathcal{C}_{1,p+1}, \mathcal{C}_{2,p+1}, \ldots, \mathcal{C}_{p-1,p+1}, \mathcal{C}_{1,p+2}, \ldots, \mathcal{C}_{p-1,p+2}, \ldots, \mathcal{C}_{1,p+q}, \ldots, \mathcal{C}_{p-1,p+q}, \\
& \mathcal{C}_{p+1,p+1}, \\
& \mathcal{C}_{p,p+1}, \mathcal{C}_{p,p+2}, \ldots, \mathcal{C}_{p,p+q}, \\
& \mathcal{C}_{p+1,p+1}, \mathcal{C}_{p+2,p+2}, \mathcal{C}_{p+1,p+2}, \mathcal{C}_{p+2,p+2}, \ldots, \mathcal{C}_{p+1,p+q}, \ldots, \mathcal{C}_{p+q-2,p+q}, \mathcal{C}_{p+q,p+q}, \\
& \mathcal{C}_{p+q-1,p+q}, \mathcal{C}_{p+q,p+q}.
\end{aligned}
$$

In Fig. A.1, we have colored those parts of the matrix in which the pivot elements, corresponding to the pivot pair sequences from the second, third and fourth line in the definition of $\mathcal{I}'_{\mathcal{A}_2}$ lie. The arrows indicate the orderings of the annihilations which are used within the blocks. The numbers in the lower triangle (we use the symmetry property of the matrix) indicate the ordering used on the macro-level.

We see that consecutive hyperbolic transformations are defined by the pivot pairs from the second and fourth line in the definition of $\mathcal{I}'_{\mathcal{A}_2}$ (numbers 1–9 and 11–13 in the figure). They are separated by the sequence $\mathcal{C}_{p+1,p+1}$ in the third line, which corresponds to orthogonal transformations (and is indicated by the number 10 in the figure). Note however, that orthogonal transformations defined by $\mathcal{C}_{p+1,p+1}$ do not change the norm of any block within $\mathbf{A}_{12}$. Therefore, we can still obtain (A.2), but the proof which leads to it has to be divided into two parts. In the first part, one shows that $\|[e_1, e_2, \ldots, e_{\nu-n_p}]^T \mathbf{A}_{12}^{(\tau M)}\|$ tends to zero as $\tau \to \infty$. In the second part one shows that $\|[e_{\nu-n_p+1}, \ldots, e_\nu]^T \mathbf{A}_{12}^{(\tau M)}\|$ tends to zero as $\tau \to \infty$, and for this part, we can introduce $\bar{k}'_\tau = k'_\tau + (\nu - n_p)(n-\nu) + n_{p+1}(n_{p+1}-1)/2$, which takes the role of k'_τ in (A.2). □

Proof of Lemma 3.5. Any strategy from the class $\mathcal{A}_1$ can be replaced by the equivalent, column-cyclic strategy (note that $M = N$). By Lemma 3.1, the column-cyclic strategy can be replaced by $\mathcal{I}_0$ which is defined by the sequence of pairs $[\mathcal{O}_\mathcal{C}(\mathbf{A}_{11}), \mathcal{O}_\mathcal{C}(\mathbf{A}_{12}), \mathcal{O}_\mathcal{C}(\mathbf{A}_{22})]$. By this strategy, all elements of the

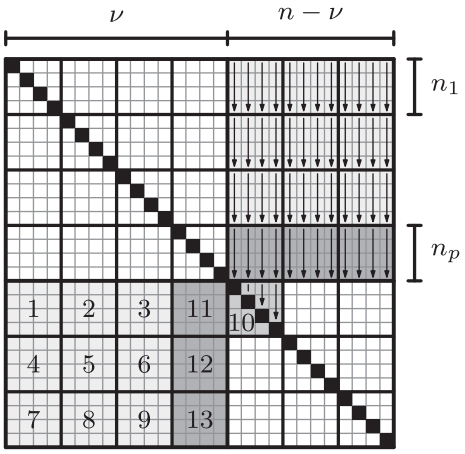

Fig. A.1. Orderings of annihilation.

upper-triangle of $\mathbf{A}_{11}$ are annihilated by columns. Hence, the result of Henrici and Zimmerman [24] implies

$$\mathrm{Off}(\mathbf{A}_{11}^{(k'_\tau)}) \leqslant \mu \mathrm{Off}\left(\mathbf{A}_{11}^{((\tau-1)M)}\right), \quad \mu = \sqrt{1 - 2^{-\frac{(\nu-1)(\nu-2)}{2}}}, \quad k'_\tau = (\tau-1)M + \frac{\nu(\nu-1)}{2}.$$

The subsequent $\nu(n-\nu)$ hyperbolic steps have a vanishing impact on $\mathrm{Off}(\mathbf{A}_{11}^{(k'_\tau)})$. Indeed, if $\mathbf{W}_\tau$ is the product of all $\nu(n-\nu)$ hyperbolic rotations within the τth sweep, then by (2.17) we have $\mathbf{W}_\tau \to I_n$ as $\tau \to \infty$. Because of (2.14) and because the last $(n-\nu)(n-\nu-1)/2$ simple rotations do not affect $\mathbf{A}_{11}^{(k'_\tau+\nu(n-\nu))}$, we can write

$$\mathrm{Off}\left(\mathbf{A}_{11}^{(\tau M)}\right) \leqslant \mu \mathrm{Off}\left(\mathbf{A}_{11}^{((\tau-1)M)}\right) + \nu_\tau, \quad 0 \leqslant \mu < 1, \quad \nu_\tau \geqslant 0, \tag{A.3}$$

where $\nu_\tau \to 0$ as $\tau \to \infty$. Now, using the second part of the proof of [17, 2.2 Lemma] (from the relation (2.6) in [17, 2.2 Lemma]), one immediately obtains $\mathrm{Off}(\mathbf{A}_{11}^{(\tau M)}) \to 0$ as $\tau \to \infty$.

Let $\mathcal{I}$ be any strategy from the class $\mathcal{A}_2$. We can replace it by the equivalent strategy $\mathcal{I}_{\mathcal{A}_2}$ which is defined above. We see that $\mathcal{I}_{\mathcal{A}_2}$ is block-column oriented, hence after

$$k'_\tau - (\tau-1)M = \frac{\nu(\nu-1)}{2} + \sum_{i=1}^{p} \frac{n_1(n_1-1)}{2}$$

simple rotations in the τ th quasi-cycle, the submatrix $\mathbf{A}_{11}^{(k'_\tau)}$ of order ν has undergone a full quasi-sweep consisting of orthogonal Jacobi transformations, under the same pivot strategy $\mathcal{I}_{\mathcal{A}_2}$. By the result from [20, Theorem 2.1], we have

$$\mathrm{Off}(\mathbf{A}_{11}^{(k'_\tau)}) \leqslant \mu \mathrm{Off}\left(\mathbf{A}_{11}^{((\tau-1)M)}\right), \quad 0 \leqslant \mu < 1,$$

where for each partition $(n_1, \ldots, n_p)$ of ν, μ depends just on ν. The rest of the proof is the same as earlier.

Let $\mathcal{I}$ be any strategy from the class $\mathcal{A}_3$. We can replace it by the equivalent strategy which is defined by the ordering Θ',

$$\Theta' = [O_2, O_1, O_3, O_2, O_1, O_5, O_4, O_6, O_5, O_4, O_7]. \tag{A.4}$$

For the strategy defined by Θ', let

$$k'_\tau = (\tau-1)M + \frac{\nu(\nu-1)}{2} + \frac{\alpha_1(\alpha_1-1)}{2} + \frac{\alpha_2(\alpha_2-1)}{2}, \tag{A.5}$$

where α_1 and α_2 are given in (2.19). Note that $\alpha_1 + \alpha_2 = \nu$. We see that within the first $k'_\tau - (\tau-1)M$ transformations with simple rotations, there exists a full cycle of $\nu(\nu-1)/2$ rotations, defined by $[O_1, O_3, O_2]$. Since $[O_1, O_3, O_2] \sim O_C(\mathbf{A}_{11})$, we can use again the result of Henrici and Zimmerman [24], to obtain

$$\mathrm{Off}\left(\mathbf{A}_{11}^{(k'_\tau)}\right) \leqslant \mu \mathrm{Off}\left(\mathbf{A}_{11}^{((\tau-1)M)}\right), \quad \mu = \sqrt{1 - 2^{-(\nu-1)(\nu-2)/2}}. \tag{A.6}$$

Here we have used the fact that orthogonal Jacobi rotation at step k cannot increase $\mathrm{Off}(\mathbf{A}_{11}^{(k)})$. The next

$$\frac{(n-\nu)(n-\nu-1)}{2} + \frac{\beta_1(\beta_1-1)}{2} + \frac{\beta_2(\beta_2-1)}{2}$$

orthogonal rotations have no impact on $\mathbf{A}_{11}^{(k'_\tau)}$ and the final $\nu(n-\nu)$ hyperbolic transformations have a vanishing impact on $\mathbf{A}_{11}^{(k'_\tau+\nu(n-\nu))}$. This means that the relation (A.6) takes form (A.3) with $\nu_\tau \to 0$ as $\tau \to \infty$. The rest of the proof is the same as above. □

Proof of Lemma 3.6. We start from an arbitrary strategy from $\mathcal{A}_1$ and replace it by the row-cyclic strategy. The row-cyclic strategy is equivalent to the cyclic strategy defined by the sequence of pairs $[O_R(\mathbf{A}_{11}), O_R(\mathbf{A}_{12}), O_R(\mathbf{A}_{22})]$. Using that strategy, let $k'_\tau = (\tau-1)M + \nu(\nu-1)/2$ and $k''_\tau = k'_\tau + \nu(n-\nu)$ with $M = N$. Obviously, $\mathbf{A}_{22}^{(k'_\tau)} = \mathbf{A}_{22}^{((r-1)M)}$ and if $\mathbf{W}_\tau$ is the product of all $\nu(n-\nu)$ hyperbolic rotations within this τth sweep, we have $\mathbf{W}_\tau \to I_n$ as $\tau \to \infty$. Therefore, we have

$$\mathrm{Off}\left(\mathbf{A}_{22}^{(k''_\tau)}\right) \leqslant \mathrm{Off}\left(\mathbf{A}_{22}^{((\tau-1)M)}\right) + \tilde{\nu}_\tau, \quad \tilde{\nu}_\tau \geqslant 0, \tag{A.7}$$

with $\tilde{\nu}_\tau \to 0$ as $\tau \to \infty$. The last $(n-\nu)(n-\nu-1)/2$ orthogonal Jacobi steps under the row-cyclic strategy will decrease the off-norm of $\mathbf{A}_{22}^{(k''_\tau)}$. Using again the result of Henrici and Zimmermann and (A.7), we obtain

$$\mathrm{Off}(\mathbf{A}_{22}^{(\tau M)}) \leqslant \mu(\mathrm{Off}(\mathbf{A}_{22}^{((\tau-1)M)}) + \tilde{\nu}_\tau) = \mu \mathrm{Off}(\mathbf{A}_{22}^{((\tau-1)M)}) + \nu_\tau, \quad 0 \leqslant \mu < 1,$$

where $\nu_\tau = \mu\tilde{\nu}_\tau \to 0$ as $\tau \to \infty$. The rest of the proof is as in the previous lemma.

The proof for an arbitrary strategy from the class $\mathcal{A}_3$ is very similar. Let Θ' and k'_τ be defined as in the relations (A.4) and (A.5) and let (see (2.19))

$$k''_\tau = k'_\tau + \frac{(n-\nu)(n-\nu-1)}{2} + \frac{\beta_1(\beta_1-1)}{2} + \frac{\beta_2(\beta_2-1)}{2}.$$

Using the equivalent strategy $\mathcal{I}_{\Theta'}$, we first note that Jacobi rotations which rotate the elements within $\mathbf{A}_{11}^{((\tau-1)M)}$ do not affect $\mathbf{A}_{22}^{((\tau-1)M)}$, hence $\mathbf{A}_{22}^{(k'_\tau)} = \mathbf{A}_{22}^{((\tau-1)M)}$. Since in the sequence of pivot pairs $[O_5, O_4, O_6, O_5, O_4]$, we recognize that $[O_4, O_6, O_5] \sim O_C(\mathbf{A}_{22})$, we have $\mathrm{Off}(\mathbf{A}_{22}^{(k''_\tau)}) \leqslant \mu \mathrm{Off}(\mathbf{A}_{22}^{(k'_\tau)})$ for some $0 \leqslant \mu < 1$. The remaining $\nu(n-\nu)$ hyperbolic rotations yield the relation

$$\mathrm{Off}(\mathbf{A}_{22}^{(\tau M)}) \leqslant \mathrm{Off}(\mathbf{A}_{22}^{(k''_\tau)}) + \nu_\tau \leqslant \mu \mathrm{Off}(\mathbf{A}_{22}^{(k'_\tau)}) + \nu_\tau = \mu \mathrm{Off}(\mathbf{A}_{22}^{((\tau-1)M)}) + \nu_\tau,$$

where $0 \leqslant \mu < 1$, $\nu_\tau \geqslant 0$, and $\nu_\tau \to 0$ as $\tau \to \infty$. As earlier, we obtain $\mathrm{Off}(\mathbf{A}_{22}^{(rM)}) \to 0$ as $r \to \infty$.

Finally, let us choose an arbitrary strategy from the class $\mathcal{A}_2$. We can replace it by the equivalent strategy $\mathcal{I}'_{\mathcal{A}_2}$ which is given above. Let

$$k'_\tau = (\tau-1)M + \tilde{k}, \qquad \tilde{k} = \frac{\nu(\nu-1)}{2} + \sum_{i=1}^{p} \frac{n_i(n_i-1)}{2} + (\nu - n_p)(n-\nu),$$

$$k''_\tau = k'_\tau + \frac{n_{p+1}(n_{p+1}-1)}{2}, \qquad k'''_\tau = k''_\tau + n_p(n-\nu).$$

Here k'_τ refers to the stage reached after the orthogonal and hyperbolic transformations, corresponding to the pairs in the first three lines of the definition of $\mathcal{I}'_{\mathcal{A}_2}$, have been applied. Similarly, k''_τ and k'''_τ correspond to the fourth and fifth line in the definition of $\mathcal{I}'_{\mathcal{A}_2}$.

As earlier, from the relations (2.14) and (2.17), we can conclude that

$$\mathbf{A}_{22}^{(k'_\tau)} = \mathbf{A}_{22}^{((r-1)M)} + G'_\tau, \quad G'_\tau \to 0 \quad \text{as} \quad \tau \to \infty. \tag{A.8}$$

Here, G'_τ is the perturbation coming from the hyperbolic transformations.

Let $\mathbf{R}_{p+1,p+1}$ denote the product of the next $n_{p+1}(n_{p+1}-1)/2$ orthogonal Jacobi rotations, defined by $O_C(A_{p+1,p+1})$. Then

$$\mathbf{A}_{22}^{(k''_\tau)} = \mathbf{R}_{p+1,p+1}^T \mathbf{A}_{22}^{((\tau-1)M)} \mathbf{R}_{p+1,p+1} + \mathbf{R}_{p+1,p+1}^T G'_\tau \mathbf{R}_{p+1,p+1}.$$

The next $k'''_\tau - k''_\tau$ hyperbolic rotations have a vanishing effect on $\mathbf{A}_{22}^{(k''_\tau)}$, hence

$$\begin{aligned}\mathbf{A}_{22}^{(k'''_\tau)} &= \mathbf{R}_{p+1,p+1}^T \mathbf{A}_{22}^{((\tau-1)M)} \mathbf{R}_{p+1,p+1} + \mathbf{R}_{p+1,p+1}^T G'_\tau \mathbf{R}_{p+1,p+1} + G''_\tau, \\ &= \mathbf{R}_{p+1,p+1}^T \mathbf{A}_{22}^{((\tau-1)M)} \mathbf{R}_{p+1,p+1} + G'''_\tau, \qquad G''_\tau \to 0 \quad \text{and} \quad G'''_\tau \to 0 \quad \text{as} \quad \tau \to \infty.\end{aligned}$$

The remaining $rM - k''_\tau$ Jacobi rotations are defined by the sequence of pivot pairs indicated by the last two lines in the definition of $\mathcal{I}'_{\mathcal{A}_2}$. Let $\widetilde{\mathbf{R}}_\tau$ denote the product of these rotation matrices. Then

$$\begin{aligned}\mathbf{A}_{22}^{(rM)} &= \widetilde{\mathbf{R}}_\tau^T \mathbf{R}_{p+1,p+1}^T \mathbf{A}_{22}^{((\tau-1)M)} \mathbf{R}_{p+1,p+1} \widetilde{\mathbf{R}}_\tau + \widetilde{\mathbf{R}}_\tau^T G'''_\tau \widetilde{\mathbf{R}}_\tau \\ &= \mathbf{R}_\tau^T \mathbf{A}_{22}^{((\tau-1)M)} \mathbf{R}_\tau + G_\tau, \qquad G_\tau = \widetilde{\mathbf{R}}_\tau^T G'''_\tau \widetilde{\mathbf{R}}_\tau \to 0 \text{ as } \tau \to \infty.\end{aligned} \tag{A.9}$$

We claim that (A.9) describes a Jacobi-type process, of the form (3.3), for $H = \mathbf{A}_{22}^{(\tilde{k})} = \mathbf{A}_{22} + G'_1$, defined by the strategy $\mathcal{I}_{\mathcal{A}_2}$, which is $\pi/4$-convergent. Here G'_1 is from the relation (A.8) with $\tau = 1$. Indeed, each quasi-sweep consists of

$$M = \frac{(n-\nu)(n-\nu-1)}{2} + \sum_{i=p+1}^{p+q} \frac{n_i(n_i-1)}{2}$$

similarity transformations with rotations. In the quasi-cycle τ, all the perturbations $F^{(k)}$ are zero except the one defined by $k = (\tau-1)M + n_{p+1}(n_{p+1}-1)/2$ (it corresponds to G''_τ in the above relation) and the one defined by $k = \tau M$ (it corresponds to G'_τ above). For this process the conditions (i)–(iii) and (a) of Lemma 3.4 are fulfilled. So, we can conclude that $\mathrm{Off}(\mathbf{A}_{22}^{(\tau M)}) \to 0$ as $\tau \to \infty$. □

References

[1] P. Arbenz, I. Slapničar, An analysis of parallel implementations of the block-Jacobi algorithm for computing the SVD, in: D. Kalpić, V. Hljuz Dobrić (Eds.), Proceedings of the 17th International Conference on Information Technology Interfaces ITI'95, Pula, Croatia, 1995.

[2] C. Ashcraft, R.G. Grimes, J.G. Lewis, Accurate symmetric indefinite linear equation solvers, SIAM J. Matrix Anal. Appl. 20 (2) (1999) 513–561.

[3] A.W. Bojanczyk, R. Onn, A.O. Steinhardt, Existence of the hyperbolic singular value decomposition, Linear Algebra Appl. 185 (1993) 21–30.

[4] J.R. Bunch, Analysis of the diagonal pivoting method, SIAM J. Numer. Anal. 8 (4) (1971) 656–680.

[5] J.R. Bunch, Partial pivoting strategies for symmetric matrices, SIAM J. Numer. Anal. 11 (3) (1974) 521–528.

[6] J.R. Bunch, L.C. Kaufman, Some stable methods for calculating inertia and solving symmetric linear systems, Math. Comp. 31 (137) (1977) 163–179.

[7] J.R. Bunch, L.C. Kaufman, B.N. Parlett, Decomposition of a symmetric matrix, Numer. Math. 27 (1) (1976) 95–109.

[8] J.R. Bunch, B.N. Parlett, Direct methods for solving symmetric indefinite systems of linear equations, SIAM J. Numer. Anal. 8 (4) (1971) 639–655.

[9] J.W. Demmel, K. Veselić, Jacobi's method is more accurate than QR, SIAM J. Matrix Anal. Appl. 13 (4) (1992) 1204–1245.

[10] F.M. Dopico, J.M. Molera, J. Moro, An orthogonal high relative accuracy algorithm for the symmetric eigenproblem, SIAM J. Matrix Anal. Appl. 25 (2) (2003) 301–351.

[11] Z. Drmač, V. Hari, On quadratic convergence bounds for the *J*-symmetric Jacobi method, Numer. Math. 64 (1) (1993) 147–180.
[12] Z. Drmač, K. Veselić, New fast and accurate Jacobi SVD algorithm: I, Tech. Rep. 169, LAPACK Working Note (Aug. 2005). URL http://www.netlib.org/lapack/lawnspdf/lawn169.pdf.
[13] Z. Drmač, K. Veselić, New fast and accurate Jacobi SVD algorithm. I, SIAM J. Matrix Anal. Appl. 29 (4) (2008) 1322–1342.
[14] Z. Drmač, K. Veselić, New fast and accurate Jacobi SVD algorithm. II, SIAM J. Matrix Anal. Appl. 29 (4) (2008) 1343–1362.
[15] E.R. Hansen, On Jacobi Methods and Block-Jacobi Methods for Computing Matrix Eigenvalues, Ph.D. thesis, Stanford University, 1960.
[16] E.R. Hansen, On cyclic Jacobi methods, SIAM J. 11 (2) (1963) 448–459.
[17] V. Hari, On the global convergence of the Eberlein method for real matrices, Numer. Math. 39 (3) (1982) 361–369.
[18] V. Hari, On some new applications of the CS decomposition, in: T.E. Simos, C. Tsitouras (Eds.), ICNAAM 2004, International Conference on Numerical Analysis and Applied Mathematics 2004, Chalkis, Greece, September 10–14, 2004, Wiley-VCH, Weinheim, 2004.
[19] V. Hari, Accelerating the SVD block-Jacobi method, Computing 75 (2005) 27–53.
[20] V. Hari, Convergence of a block-oriented quasi-cyclic Jacobi method, SIAM J. Matrix Anal. Appl. 29 (2) (2007) 349–369.
[21] V. Hari, Quadratic convergence of a special quasi-cyclic Jacobi method, Ann. Univ. Ferrara Sez. VII Sci. Mat. 53 (2) (2007) 255–269.
[22] V. Hari, S. Singer, S. Singer, Efficient eigenvalue computation by block modification of the indefinite one-sided Jacobi algorithm, in: T.E. Simos, G. Psihoyios, C. Tsitouras (Eds.), ICNAAM 2005, International Conference on Numerical Analysis and Applied Mathematics 2005, Rhodes, Greece, September 16–20, 2005, Wiley-VCH, Weinheim, 2005.
[23] V. Hari, S. Singer, S. Singer, On full block *J*-Jacobi methods, Preprint, University of Zagreb, SIAM, J. Matrix Anal. Appl., submitted for publication.
[24] P. Henrici, K. Zimmerman, An estimate for the norms of certain cyclic Jacobi operators, Linear Algebra Appl. 1 (1968) 489–501.
[25] I.C.F. Ipsen, Relative perturbation results for matrix eigenvalues and singular values, Acta Numerica 1998, Cambridge University Press, Cambridge, 1998, pp. 151–201.
[26] W.F. Mascarenhas, On the convergence of the Jacobi method for arbitrary orderings, SIAM J. Matrix Anal. Appl. 16 (4) (1995) 1197–1209.
[27] J. Matejaš, V. Hari, Quadratic convergence estimate of scaled iterates by *J*-symmetric Jacobi method, Linear Algebra Appl. 417 (2006) 434–465.
[28] G. Okša, M. Vajteršic, Efficient pre-processing in the parallel block-Jacobi SVD algorithm, Parallel Comput. 32 (2) (2006) 166–176.
[29] N.H. Rhee, V. Hari, On the global and cubic convergence of a quasi-cyclic Jacobi method, Numer. Math. 66 (1) (1993) 97–122.
[30] G. Shroff, R.S. Schreiber, On the convergence of the cyclic Jacobi method for parallel block orderings, SIAM J. Matrix Anal. Appl. 10 (3) (1989) 326–346.
[31] S. Singer, Indefinite QR factorization, BIT 46 (1) (2006) 141–161.
[32] I. Slapničar, Highly accurate symmetric eigenvalue decomposition and hyperbolic SVD, Linear Algebra Appl. 358 (2003) 387–424.
[33] K. Veselić, A Jacobi eigenreduction algorithm for definite matrix pairs, Numer. Math. 64 (1) (1993) 241–269.